Bone, Union, and Osteoarthritis

Judith Annika Willems

Bone, Union, and Osteoarthritis

Judith Annika Willems

ISBN: 978-94-6483-755-1
Cover design: Joey Roberts | www.ridderprint.nl
Lay-out and printing: Ridderprint | www.ridderprint.nl

Bone, Union, and Osteoarthritis

Bot, consolidatie en artrose

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus

Prof. dr. A.L. Bredenoord

en volgens besluit van het college voor promoties.
De openbare verdediging zal plaatsvinden op

woensdag 10 april 2024 om 15.30 uur

door

Judith Annika Willems
geboren te Rotterdam

Erasmus University Rotterdam

Erasmus

Promotiecommissie

Promotor: Prof. dr. S.M.A. Bierma-Zeinstra

Overige leden: Prof. dr. M.C. Zillikens
Prof. dr. M.H.J. Verhofstad
Prof. dr. J.L. Tol

Copromotoren: Dr. D.E. Meuffels
Dr. O.P. van der Jagt

Contents

Chapter 1	General introduction	7
Chapter 2	Accuracy of performed correction of open versus closed wedge high tibial osteotomies, with locking plate fixation: A randomized controlled trial	19
Chapter 3	Talocrural arthrodesis increases osteoarthritis severity in adjacent joints; a midterm computed tomography follow-up study	39
Chapter 4	Bone union assessment with computed tomography (CT) and statistical associations with mechanical or histological testing: A systematic review of animal studies.	55
Chapter 5	How to assess consolidation after foot and ankle arthrodesis with computed tomography. A systematic review.	79
Chapter 6	Extracorporeal shock wave treatment for delayed-union and nonunion fractures; a systematic review	105
Chapter 7	Extracorporeal shock wave therapy (ESWT) for talocrural arthrodesis; a double-blinded randomized controlled trial	127
Chapter 8	General discussion	145
Chapter 9	Summary	163
Appendices	Nederlandse samenvatting	170
	Dankwoord	176
	Curriculum vitae	180
	PhD portfolio	182
	List of publications	184

Chapter 1

General introduction

This thesis covers these three themes, osteoarthritis (OA), bone union assessment, and extracorporeal shock wave therapy (ESWT). The themes are combined and culminated in chapter 7; extracorporeal shock wave therapy compared to sham-ESWT per operatively for talocrural arthrodesis; a double-blinded randomized controlled trial (RCT).

Part 1: Osteoarthritis

OA is a complex chronic joint disease which occurs in 80% of the population aged over 70 years[1]. In 2021 more than 1.5 million people in the Netherlands had OA[2]. Joint pain is the most common complaint in OA patients, but many patients also experience other symptoms like stiffness, reduced range of motion, crepitus, joint instability, swelling and muscle weakness[1, 3]. These complaints can lead to severe impairment of physical function and decreased quality of life[4]. In the next decades, the prevalence and burden of disease of OA are expected to increase enormously due to aging of the population and rising obesity rates[3].

OA can occur in every joint of the human body. In this thesis we will focus on OA in the lower extremity, specifically the knee and talocrural joint. Of all joints, the knee is the most frequently affected by OA[2, 3, 5]. In 2021, 43.700 people were diagnosed with knee OA by their general practitioner in the Netherlands, bringing the total number of patients with knee OA to 762.700[2]. Risk factors for development of knee OA are obesity, female gender and previous knee injury[6].

Talocrural OA is much less common than knee OA and is less frequently studied[4]. The prevalence of OA in the talocrural joint has been estimated to be between 3 and 6% in adults aged 50 years and older[4, 7, 8]. It is often caused by trauma or inflammatory abnormalities[1, 4, 5]. Other risk factors for talocrural OA are smoking, obesity and hemophilia[4].

The pathogenesis of OA is a widely studied subject. In the past, it was thought that osteoarthritis was caused by wear of cartilage in the joint, without a physiological response to repair this damage[1]. However, recent insights have shown that cartilage degeneration is not the only factor involved in OA. OA is a multifactorial, whole joint disease in which multiple tissues are affected. Changes that occur during OA are: decrease of articular cartilage thickness, increase of subchondral bone thickness and an inflammation of the synovial membrane[1, 3, 9]. Also other structures are affected like the joint capsule, ligaments and periarticular muscles[3]. It is generally assumed that OA starts with a lesion in the cartilage of a joint, caused by excessive loading or abnormal shape of the joint[1]. It has been hypothesized that cartilage is already more vulnerable

to lesions due to compositional changes of the cartilage[3]. Attempts to repair cartilage lesions mainly by the release of inflammatory mediators fail, also because synovial tissue also starts to produce pro-inflammatory factors[3, 10]. The-over expression of inflammatory mediators from chondrocytes and synovial tissue probably causes further joint destruction[1, 3].

At this moment no curing treatment exists for osteoarthritis[1]. Several pharmacological treatments are currently studied to slow the process of osteoarthritis or cure osteoarthritis, but no clinical applicable treatment is available yet[3].

In patients with mild and moderate osteoarthritis, conservative treatments are available to slow the progress of OA and to improve function and reduce pain. In knee OA several conservative treatments have been shown to be effective[11]. Exercise therapy aiming to increase muscle strength, improving aerobic fitness, and reducing body weight has been shown to decrease pain and increase range of motion[3, 11]. Oral pharmacological treatments such as paracetamol and non-steroidal anti-inflammatory drugs (NSAIDs) can also be used to reduce pain. Also, topical NSAID's are strongly recommended for patients with knee OA[11]. If medication is not effective, intra-articular injections with corticosteroids can be applied in specific cases[3, 11].

Conservative talocrural OA treatments are far less studied than knee OA treatments[4]. There is no consensus guideline for adequate conservative treatments. Although evidence is limited, it seems that physical therapy may reduce pain in talocrural OA[4]. The use of intra-articular hyaluronic acid injections and oral NSAIDs can be considered but high quality studies are needed to prove effectiveness of these interventions. Also, shoe adaptations can be considered, but there is no strong scientific evidences that prove effectiveness[4]. It is also unknown if weight loss is an adequate intervention for talocrural OA[4].

In case of failure of conservative OA treatments, surgical options can be considered. The most common surgical interventions for OA are joint replacement, joint distraction, osteotomy, or arthrodesis. Which surgery is most adequate depends on which joint is affected, patient characteristics, and severity of OA. In this thesis we focus on two surgical interventions: osteotomy for patients with moderate knee OA and arthrodesis for patients with end-stage talocrural OA.

An osteotomy of the knee is a surgery for patients with moderate unicompartmental knee OA and knee malalignment. In case of medial knee OA and varus malalignment, the malalignment is corrected to a slight valgus alignment, with a high tibia osteotomy

(HTO). Herewith the load is transferred from the medial compartment to the healthy lateral compartment. Due to this shift in loading, 8 out of 10 patients can postpone a total knee arthroplasty by more than 10 years[1, 12]. Different techniques do exist to perform a HTO. The most common techniques are the lateral closed wedge technique and the medial open wedge technique. Although both techniques have advantages and disadvantages, it is unclear which technique is the best[13].

A talocrural arthrodesis is a surgical intervention for patients with end-stage talocrural OA. During this surgical intervention, remaining cartilage is removed and the joint is fixated with screws. After the surgery the tibia and talus are supposed to fuse. Although movement in the talocrural joint is impossible after talocrural arthrodesis, overall function increases as a result of decreased disability and pain reduction[14]. However, concerns have been rising about the long-term effects of talocrural arthrodesis. It has been shown that patients with talocrural arthrodesis frequently suffer from OA in the adjacent joints (talocrural, calcaneocuboid and subtalar joints)[15]. It has been hypothesized that this may be caused by altered biomechanics, in which adjacent joints compensate for the loss of motion in the talocrural joint[15]. However, based on the current literature it is unclear whether adjacent joint OA is pre-existing to talocrural arthrodesis or whether it develops as a result of talocrural arthrodesis[15].

Part 2: Bone union assessment

Bone union is a frequently studied subject within orthopedics. Successful bone growth and healing is important for many orthopedic interventions, like fracture treatment, osteotomy or arthrodesis. Fracture healing can occur through two different mechanisms: endosteal bone healing (primary bone healing) and periosteal bone healing (secondary fracture healing). Primary bone healing occurs without formation of callus and through activation of osteoclasts and osteoblasts. In this type of bone healing, absolute stability is necessary. If absolute stability is not present, as for example in casting, secondary bone healing occurs. Secondary bone healing occurs with the formation of callus and subsequent bone remodeling [1].

Sometimes bone union takes longer than expected. A fracture that takes longer to heal than expected is called a delayed union. A delayed union may heal but can also evolve into nonunion, which is a state where spontaneous bone healing will not occur. There is no general consensus on the definition of nonunion but a commonly used definition is that the fracture persists for a least nine months without signs of healing for three months[16]. Several risk factors do exist for nonunion fractures such as smoking, diabetes and fracture type[16]. Nonunions can be classified into hypertrophic or atrophic

nonunions. In hypertrophic nonunions the stability of the fracture is insufficient, causing sclerosing of the fracture ends. In anthropic nonunions, bone vitality is insufficient at the fracture ends due to lack of blood supply[1, 16].

In daily practice, bone union is assessed based on a clinical assessment and on radiographic images. Clinical indications for incomplete bone union are pain during palpation, pain during axial loading and edema[1]. However, clinical assessment of bone union is a relative subjective assessment. No consensus exists regarding a valid and reliable method for objective bone union assessment[17, 18]. Consensus on bone union assessment may be particularly important in clinical studies in which bone union is the primary outcome measure. Corrales et al. (2008) assessed the definitions used for fracture healing in studies of long-bone fractures. In the 123 studies included in this review, twelve different criteria were used to clinically assess fracture union and eleven different criteria for radiographic fracture union[17]. The different criteria for bone union increase the heterogeneity between studies and make it harder to compare results between studies. Therefore, before starting a clinical study with bone union as primary outcome measure, it is important to carefully consider criteria for bone union. We should strive for consensus on one valid and reliable method for bone union assessment.

Which method is best for radiological bone union assessment may depend on the type of bone defect and the location. For instance, primary bone healing without callus formation may be assessed differently than secondary bone healing with callus formation. Also, sometimes bone union cannot be accurately assessed from plain radiographs. For example, it has been shown that bone union after talocrural arthrodesis cannot be accurately assessed from radiographs as they lack bony details[19]. Therefore, computed tomography (CT) is advised for bone union assessment after talocrural arthrodesis[19, 20]. In this thesis we further investigated objective and valid assessment of bone union with CT, specifically after talocrural arthrodesis.

Part 3: Extracorporeal shock wave therapy (ESWT)

Extracorporeal shock wave therapy is a noninvasive therapy used in urology for the treatment of kidney stones. However, it has been shown that ESWT may also be effective to treat bone-healing problems. Around 1990, a German urologist described cortical changes in the iliac bone in patients treated for kidney stones with ESWT[21]. To further investigate this, the urologist prof. dr. Haupt, performed a pre-clinical study on humeri fracture healing in rats. In this study, significant better radiological fracture healing was seen in humeri treated with ESWT compared to sham-treated control rats[22]. In the same time, Valchanou and Michailov published a clinical study about treatment of

delayed and nonunion fractures. In this study, 82 bones with delayed and nonunion fractures were treated with ESWT. Radiological examination showed that fracture healing occurred in 85.4% of the fractures[23]. Ever since several pre-clinical and clinical studies have been performed to further investigate the effect of ESWT on bone healing[21, 24]. A major advantage of ESWT, compared to surgical intervention, is that this is a noninvasive treatment and therefor there seems to be no risk of major complications[24]. Nowadays, ESWT is the first choice treatment for nonunion fractures in Austria[21].

During ESWT acoustic pressure waves are generated. The pressure waves that are generated during ESWT are characterized by a rapid pressure increase, leading to a high peak pressure (up to 100 MPa) followed by a negative tensile wave[25]. ESWT pressure waves can be generated by three different mechanisms, namely electrohydraulic, piezoelectric or electromagnetic[26]. All three techniques generate a shock wave in a fluid medium, like water. For instance, an electrohydraulic shockwave generator contains a treatment head in which an electrode is placed and which is filled with water. A spark is created between the tips of the electrode. The spark heats the surrounding water, creating gas bubbles filled with vapor water. These bubbles expand causing a positive pressure wave. The subsequent implosion of the bubbles causes the negative pressure wave[25]. The pressure waves generated by the electrode expand spherically through the water. Due to the ellipsoid formed treatment head the pressure waves are reflected and focused on one focal point. This focal point should be targeted at the tissue to be treated, for instance a fracture. To transfer the pressure waves from the water to the human body, a contact medium, like ultrasonic gel, is used[25, 26]. The above-described ESWT is called focused ESWT. Another type of ESWT is radial ESWT. Shock waves in radial ESWT are generated through a different mechanism and are unfocused[26]. However, in this thesis we will concentrate on focused ESWT.

Basic research has shown that ESWT stimulates the release of several angiogenetic and osteogenic growth factors [27-32]. However, the exact working mechanism is still unclear. It has been hypothesized that the high peak pressure and rapid changes in pressure cause compressive, shear and tensile forces within bony tissue cells[21]. The deformation of cells as a result of these pressure changes may trigger biochemical responses. This process, during which biomechanical impulses are translated to biochemical responses is called mechanotransduction[33].

Although ESWT seems to be a promising treatment for patients with bone healing problems, it is not a globally used treatment yet. For ESWT to become first choice treatment for nonunion fractures, the effectiveness must be proven. Therefore, we should

start by generating an overview of the currently available literature on the effectiveness of ESWT for delayed- and nonunion fractures.

Where it all comes together….

This introduction covered three themes, which are each important for this final part. As mentioned before, talocrural arthrodesis is a widely accepted treatment for end-stage talocrural osteoarthritis. However, a feared complication after talocrural arthrodesis are bone union problems. Nonunion rates after talocrural arthrodesis have been reported to be around 10%[34]. Nonunions are usually treated with a revision surgery during which bone grafts are used to stimulate union. Solid union after arthrodesis is important because patients with nonunions have been shown to have poorer functional outcomes compared to fused arthrodesis and persistent pain[35]. Also, treatment of nonunions has been shown to be expensive[36]. Therefore, it would be interesting if we could find a way to stimulate union after talocrural arthrodesis. This brings us to ESWT. ESWT may be an effective treatment to decrease numbers of delayed- and nonunions after talocrural arthrodesis. To study the effectiveness of ESWT after talocrural arthrodesis, a valid and reliable method must be used to assess bone union after talocrural arthrodesis. Therefore, this subject was extensively studied within this thesis.

Aims and outline of the thesis

Part 1

In the first study of the thesis, which was a RCT, we compared two techniques to perform HTO: the medial open wedge technique and the lateral closed wedge technique. We focused on the accuracy of correction of both techniques, and investigated which technique is most accurate in reaching the planned correction. The findings of this randomized controlled trial are presented in **chapter 2**.

In the next chapter we focus on patients with severe talocrural OA. Severe talocrural OA can be treated by performing a talocrural arthrodesis. A possible complication of this treatment is the development of OA in adjacent joints. However, although OA is found in adjacent joints after talocrural arthrodesis, it is currently unclear whether this OA is already present at the time of the talocrural arthrodesis, or whether it develops afterwards. We performed a mid-term follow-up study with pre-operative CT and follow-up CT to accurately assess the degree of preoperative and follow-up OA in adjacent joints. This study is presented in **chapter 3**.

Part 2

In **chapter 4 and 5** we focus on the methodology of bone union assessment. In clinical practice, bone union is assessed based on physical examination and radiological images (radiographs or CT-scan). However, the assessment of bone union is quite subjective. For scientific and clinical purposes it would be valuable if an accurate and valid method for bone union assessment would exist. We therefore performed a systematic review in pre-clinical studies to investigate which CT-parameters correlate with actual bone union (measured by histological or biomechanical testing). This systematic review is presented in **chapter 5**. In **chapter 6** we performed a systematic review in which we investigated currently used methods for bone union assessment with CT after talocrural arthrodesis. In this review we formulated an advise on how bone union should be assessed after foot and talocrural arthrodesis.

Part 3

ESWT seems to be a promising therapy to stimulate bone union. To get more insights in the available literature on the effect of ESWT on bone union, we conducted a systematic review. We focused on the effects of ESWT in delayed-healing and nonunion fractures. This systematic review is presented in **chapter 6** of this thesis.

Where it all comes together…..

In **chapter 7** a randomized controlled trial is presented that investigated whether ESWT is an effective treatment to reduce the number of delayed unions after talocrural arthrodesis.

Discussion and summary

Chapter 8 contains a general discussion on the findings of this thesis. Also, the limitations of this thesis will be discussed and future perspectives will be highlighted.

Chapter 9 provides a summary of this thesis.

References

[1] J. A. N. v. M. Verhaar, J.B.A., Orthopaedics and Traumatology The Essential Guide. Bohn stafleu van loghum 2021.

[2] "Artrose." Rijksinstituut voor Volksgezondheid en Milieu. https://www.vzinfo.nl/artrose#node-prevalentie-en-aantal-nieuwe-gevallen-van-artrose (accessed 2022-11-11.

[3] D. J. Hunter and S. Bierma-Zeinstra, "Osteoarthritis," The Lancet, vol. 393, no. 10182, pp. 1745-1759, 2019/04/27/ 2019, doi: https://doi.org/10.1016/S0140-6736(19)30417-9.

[4] J. B. Arnold et al., "International Foot and Ankle Osteoarthritis Consortium review and research agenda for diagnosis, epidemiology, burden, outcome assessment and treatment," Osteoarthritis and Cartilage, vol. 30, no. 7, pp. 945-955, 2022/07/01/ 2022, doi: https://doi.org/10.1016/j.joca.2022.02.603.

[5] S. Swain et al., "Trends in incidence and prevalence of osteoarthritis in the United Kingdom: findings from the Clinical Practice Research Datalink (CPRD)," Osteoarthritis Cartilage, vol. 28, no. 6, pp. 792-801, Jun 2020.

[6] V. Silverwood, M. Blagojevic-Bucknall, C. Jinks, J. L. Jordan, J. Protheroe, and K. P. Jordan, "Current evidence on risk factors for knee osteoarthritis in older adults: a systematic review and meta-analysis," Osteoarthritis Cartilage, vol. 23, no. 4, pp. 507-15, Apr 2015.

[7] E. Roddy et al., "The population prevalence of symptomatic radiographic foot osteoarthritis in community-dwelling older adults: cross-sectional findings from the Clinical Assessment Study of the Foot," Annals of the Rheumatic Diseases, vol. 74, no. 1, p. 156, 2015, doi: 10.1136/annrheumdis-2013-203804.

[8] C. Murray, M. Marshall, T. Rathod, C. J. Bowen, H. B. Menz, and E. Roddy, "Population prevalence and distribution of ankle pain and symptomatic radiographic ankle osteoarthritis in community dwelling older adults: A systematic review and cross-sectional study," PLoS One, vol. 13, no. 4, p. e0193662, 2018.

[9] K. D. Brandt, E. L. Radin, P. A. Dieppe, and L. Van De Putte, "Yet more evidence that osteoarthritis is not a cartilage disease," vol. 65, ed: BMJ Publishing Group Ltd, 2006, pp. 1261-1264.

[10] M. Rahmati, A. Mobasheri, and M. Mozafari, "Inflammatory mediators in osteoarthritis: A critical review of the state-of-the-art, current prospects, and future challenges," Bone, vol. 85, pp. 81-90, 2016/04/01/ 2016, doi: https://doi.org/10.1016/j.bone.2016.01.019.

[11] R. R. Bannuru et al., "OARSI guidelines for the non-surgical management of knee, hip, and polyarticular osteoarthritis," Osteoarthritis and Cartilage, vol. 27, no. 11, pp. 1578-1589, 2019/11/01/ 2019, doi: https://doi.org/10.1016/j.joca.2019.06.011.

[12] C. A. Primeau et al., "Total knee replacement after high tibial osteotomy: time-to-event analysis and predictors," Cmaj, vol. 193, no. 5, pp. E158-E166, Feb 1 2021.

[13] R. W. Brouwer et al., "Osteotomy for treating knee osteoarthritis," Cochrane Database Syst Rev, vol. 2014, no. 12, p. CD004019, Dec 13 2014.

[14] M. Glazebrook, B. N. Burgesson, A. S. Younger, and T. R. Daniels, "Clinical outcome results of total ankle replacement and ankle arthrodesis: a pilot randomised controlled trial," Foot Ankle Surg, vol. 27, no. 3, pp. 326-331, Apr 2021.

[15] J. S. Ling et al., "Investigating the relationship between ankle arthrodesis and adjacent-joint arthritis in the hindfoot: a systematic review," J Bone Joint Surg Am, vol. 97, no. 6, pp. 513-20, Mar 18 2015.

[16] J. D. K. Thomas, J.L. . "Bone Nonunion." https://www.ncbi.nlm.nih.gov/books/NBK554385/ (accessed 18-11-2022, 2022).

[17] L. A. Corrales, S. Morshed, M. Bhandari, and T. Miclau, 3rd, "Variability in the assessment of fracture-healing in orthopaedic trauma studies," J Bone Joint Surg Am, vol. 90, no. 9, pp. 1862-8, Sep 2008.

[18] S. Morshed, L. Corrales, H. Genant, and T. Miclau, 3rd, "Outcome assessment in clinical trials of fracture-healing," J Bone Joint Surg Am, vol. 90 Suppl 1, pp. 62-7, Feb 2008.

[19] M. J. Coughlin, J. S. Grimes, P. D. Traughber, and C. P. Jones, "Comparison of radiographs and CT scans in the prospective evaluation of the fusion of hindfoot arthrodesis," Foot Ankle Int, vol. 27, no. 10, pp. 780-7, Oct 2006.

[20] C. P. Jones, M. J. Coughlin, and P. S. Shurnas, "Prospective CT scan evaluation of hindfoot nonunions treated with revision surgery and low-intensity ultrasound stimulation," Foot Ankle Int, vol. 27, no. 4, pp. 229-35, Apr 2006.

[21] R. Mittermayr, N. Haffner, X. Feichtinger, and W. Schaden, "The role of shockwaves in the enhancement of bone repair - from basic principles to clinical application," Injury, vol. 52 Suppl 2, pp. S84-S90, Jun 2021.

[22] G. Haupt, A. Haupt, A. Ekkernkamp, B. Gerety, and M. Chvapil, "Influence of shock waves on fracture healing," Urology, vol. 39, no. 6, pp. 529-532, 1992/06/01/ 1992, doi: https://doi.org/10.1016/0090-4295(92)90009-L.

[23] V. D. Valchanou and P. Michailov, "High energy shock waves in the treatment of delayed and nonunion of fractures," International Orthopaedics, vol. 15, no. 3, pp. 181-184, 1991/09/01 1991, doi: 10.1007/bf00192289.

[24] W. Schaden, R. Mittermayr, N. Haffner, D. Smolen, L. Gerdesmeyer, and C. J. Wang, "Extracorporeal shockwave therapy (ESWT)--First choice treatment of fracture non-unions?," Int J Surg, vol. 24, no. Pt B, pp. 179-83, Dec 2015.

[25] J. A. Ogden, A. Tóth-Kischkat, and R. Schultheiss, "Principles of shock wave therapy," Clin Orthop Relat Res, no. 387, pp. 8-17, Jun 2001.

[26] M. C. d'Agostino, K. Craig, E. Tibalt, and S. Respizzi, "Shock wave as biological therapeutic tool: From mechanical stimulation to recovery and healing, through mechanotransduction," International Journal of Surgery, vol. 24, pp. 147-153, 2015/12/01/ 2015, doi: https://doi.org/10.1016/j.ijsu.2015.11.030.

[27] D. Sun et al., "Shockwaves Induce Osteogenic Differentiation of Human Mesenchymal Stem Cells Through ATP Release and Activation of P2X7 Receptors," Stem Cells, vol. 31, no. 6, pp. 1170-1180, 2013, doi: 10.1002/stem.1356.

[28] J.-k. Xu et al., "Optimal Intensity Shock Wave Promotes the Adhesion and Migration of Rat Osteoblasts via Integrin β1-mediated Expression of Phosphorylated Focal Adhesion Kinase," Journal of Biological Chemistry, vol. 287, no. 31, pp. 26200-26212, 2012/07/27/ 2012, doi: https://doi.org/10.1074/jbc.M112.349811.

[29] F.-S. Wang et al., "Ras Induction of Superoxide Activates ERK-dependent Angiogenic Transcription Factor HIF-1α and VEGF-A Expression in Shock Wave-stimulated Osteoblasts*," Journal of Biological Chemistry, vol. 279, no. 11, pp. 10331-10337, 2004/03/12/ 2004, doi: https://doi.org/10.1074/jbc.M308013200.

[30] Y.-J. Chen et al., "Activation of extracellular signal-regulated kinase (ERK) and p38 kinase in shock wave-promoted bone formation of segmental defect in rats," Bone, vol. 34, no. 3, pp. 466-477, 2004/03/01/ 2004, doi: https://doi.org/10.1016/j.bone.2003.11.013.

[31] C. H. Ha, S. Kim, J. Chung, S. H. An, and K. Kwon, "Extracorporeal shock wave stimulates expression of the angiogenic genes via mechanosensory complex in endothelial cells: Mimetic effect of fluid shear stress in endothelial cells," International Journal of Cardiology, vol. 168, no. 4, pp. 4168-4177, 2013/10/09/ 2013, doi: https://doi.org/10.1016/j.ijcard.2013.07.112.

[32] F. S. Wang et al., "Temporal and spatial expression of bone morphogenetic proteins in extracorporeal shock wave-promoted healing of segmental defect," Bone, vol. 32, no. 4, pp. 387-396, 2003/04/01/ 2003, doi: https://doi.org/10.1016/S8756-3282(03)00029-2.

[33] D. E. Ingber, "Cellular mechanotransduction: putting all the pieces together again," Faseb J, vol. 20, no. 7, pp. 811-27, May 2006.

[34] A. Nihal, R. E. Gellman, J. M. Embil, and E. Trepman, "Ankle arthrodesis," Foot and Ankle Surgery, vol. 14, no. 1, pp. 1-10, 2008/01/01/ 2008, doi: https://doi.org/10.1016/j.fas.2007.08.004.

[35] F. Krause et al., "Clinical Outcomes of Nonunions of Hindfoot and Ankle Fusions," J Bone Joint Surg Am, vol. 98, no. 23, pp. 20 06-2016, Dec 7 2016.

[36] O. J. Gagne, A. N. Veljkovic, M. Glazebrook, M. Penner, K. Wing, and A. S. E. Younger, "Agonizing and Expensive: A Review of Institutional Costs of Ankle Fusion Nonunions," Orthopedics, vol. 43, no. 4, pp. e219-e224, Jul 1 2020.

Chapter 2

Accuracy of performed correction of open versus closed wedge high tibial osteotomies, with locking plate fixation: A randomized controlled trial

A. Willems, T. Duivenvoorden, E.M. van Es, D.E. Meuffels, S.M.A. Bierma-Zeinstra, P.K. Bos, M. Reijman
Submitted

Abstract

Aims: The primary aim was to investigate the accuracy of performed correction between open wedge high tibia osteotomies (HTO) and closed wedge HTO. The open wedge HTO allows fine adjustment before final fixation and was therefore expected to achieve a more accurate correction compared to closed wedge HTO.

Methods: This randomized controlled trial included patients who underwent HTO. Inclusion criteria were: medial knee pain for at least three months, severity of knee pain more than 20 mm on a VAS score (range 0 to 100 mm), Kellgren & Lawrence grade 1-3, and presence of varus malalignment measured on whole leg radiograph. Patients were randomized between open or closed wedge HTO. The primary outcome was the difference in accuracy of performed correction between both groups. Accuracy of performed correction was calculated by subtracting the planned correction from the achieved correction, based on Hip-Knee-Ankle angles. Secondary outcomes were differences in loss of correction and differences in complications between the groups.

Results: Seventy-six patients were included of which 38 patients were randomized to the open wedge HTO and 38 patients to the closed wedge HTO. The accuracy of performed correction in the open wedge HTO was an under-correction of 0.9° (95% confidence interval (CI) -2.0 to 0.1) and in the closed wedge HTO an over-correction of 1.0° (95% CI -0.1 to 2.0). The difference of 1.9° was significant between the groups (95% CI 0.7 to 3.1). There were no significant differences for loss of correction and complications between the groups.

Conclusion: The accuracy of performed correction is different between open wedge HTO and closed wedge HTO. The open wedge HTO tends to under-correction the planned correction angle, whereas the closed wedge HTO slightly over-corrects. Based on the results of this study both techniques can be used to perform HTO to achieve accurate correction.

- This study showed that the open wedge HTO and closed wedge HTO can both achieve accurate corrections in patients with medial knee osteoarthritis and varus malalignment.

Introduction

Although several studies have been published about high tibial osteotomies (HTO), there is no high-level evidence as for which surgical technique is best. An HTO is a surgical intervention for patients with medial osteoarthritis (OA) of the knee and varus malalignment in which conservative treatment is unsuccessful. HTO is of particular interest for young (<65 years of age) active patients, as these patients are generally too young for a total knee arthroplasty due to the relatively limited survival of the prosthesis[1].

In patients with medial knee OA and varus alignment, the varus alignment leads to increased loading of the medial compartment, which increases the risk of medial OA progression[2]. HTO corrects the varus malalignment into valgus alignment leading to unloading of the medial affected compartment. Unloading of the medial compartment can lead to substantial reduction of pain and improvement of functional outcome[3].

HTO can be performed with an open or closed wedge HTO. The closed wedge HTO is an older technique and is assumed to be more stable[3]. The open method gained popularity due to its less invasive nature, with less risk of peroneal nerve damage and disruption to the tibiofibular joint. Also, fine adjustments before fixation are easier to perform[4]. There is no consensus as to which is better, which is confirmed in recent systematic reviews [5-8].

The open wedge HTO has the theoretical advantage of allowing fine adjustment of the correction before final fixation. When this is combined with a rigid fixation plate, we expect this procedure to achieve a more accurate and enduring correction, as compared to closed wedge HTO. The primary aim of this study was to evaluate if an open wedge HTO achieves a more accurate correction 6 weeks after HTO, compared to a closed HTO.

Method

Study design

This open-labeled randomized controlled trial (RCT) with a superiority design evaluated the accuracy of performed correction of two techniques for HTO. Patients were recruited between July 2012 and July 2020 at the Erasmus MC, University Medical Center, Rotterdam, the Netherlands. The study was approved by the local ethics committee (MEC-2010-181). The study protocol was prospectively registered at 'the International Clinical Trial Registry Platform' (https://trialsearch.who.int/; registration number: NTR3506)

Patients

Patients were recruited from the outpatient clinic of the department of Orthopedics. Inclusion criteria were: knee pain located over the medial tibiofemoral compartment for more than 3 months, severity of knee pain of more than 20 millimeters (mm) on a VAS score (with 0 mm indicating no pain and 100 mm indicating worst pain), radiographic signs of OA (Kellgren & Lawrence grade 1-3), and presence of varus malalignment (>0° and ≤12° varus) as measured on loaded whole leg radiograph. Reasons for exclusion were: concomitant radiological OA of the lateral compartment, grade-3 collateral ligament laxity, range of motion < 100°, flexion contracture of > 10°, history of fracture or previous open operation of the lower limb, history of anterior cruciate ligament reconstruction, rheumatoid arthritis, contralateral HTO that was already included in this trial, patients who were unable to attend follow-up measurements or had insufficient command of the Dutch language. Eligible patients received standardized written and oral information about the study. Patients who were interested in participating were invited for a standardized whole leg radiograph to confirm varus malalignment. Afterwards, when varus malalignment was confirmed, informed consent was signed. Baseline measurements were performed including the inventory of baseline characteristics (age, sex, height, weight, duration of complaints, number of days with pain during the last month). Also, the severity of knee OA was assessed on digitized standard weight-bearing knee posterior-anterior and lateral radiographs.

Randomization

Randomization was done after written informed consent was signed. Randomization was done by an independent person with a computer-generated randomization list. The list was built through block randomization, with variable sizes of the blocks (range 2-6), and was stratified for the orthopedic surgeon. The orthopedic surgeon and patient were informed about randomization outcome. The statistician was blinded for randomization outcomes.

Interventions

Patients were randomized to medial open wedge HTO or lateral closed wedge HTO. Preoperative planning was conducted with whole leg radiographs endeavoring to achieve an over-correction of 3 to 4° of valgus, based on the preference of the orthopedic surgeon. The possible variation of 3 to 4 degrees was necessary as the closed wedge HTO osteotomy saw guide works with two-degree increments.

Medial open wedge HTO

A medial approach was performed. The pes anserinus and the superficial medial collateral ligament were released and moved dorsally. A Kirschner wire was inserted just proximal of the tuberosity towards the tibiofibular joint under radiographic guidance. A second wire was inserted parallel to the slope direction of the previous Kirschner wire. The osteotomy was performed along these 2 Kirschner wires under X-ray control. The lateral cortex was left intact. The osteotomy was opened gradually and to the preplanned angle. The medial TomoFix™ locking compression plate (Synthes, Westchester, Pennsylvania, USA) was fixed to the proximal section of the osteotomy. The lateralization of the mechanical axis was checked with the image intensifier using a metal rod to mimic the mechanical axis (centered on the femoral head and the center of the talus, with the leg in neutral rotation) and was corrected if necessary. Then the distal section was fixed.

Lateral closed wedge HTO

An anterolateral approach was performed. An osteotomy and resection were performed through the tibiofibular joint. A Kirschner wire was placed two centimeters below and parallel to the joint line. The proximal osteotomy was performed while preserving the medial cortex. The second, and distal, osteotomy was performed just proximal of the tuberosity, using the HTO osteotomy saw guide. The bone wedge was removed. The resulting open wedge was closed. The lateralization of the mechanical axis was checked with the image intensifier using a metal rod to mimic the mechanical axis. The osteotomy was corrected if necessary. It was then fixed with the lateral TomoFix™ locking compression plate (Synthes, Westchester, Pennsylvania, USA).

Postoperative management was identical for both techniques. The day after the operation patients were mobilized with partial weight bearing (up to 50%) of the operated leg, depending on pain level. Flexion and extension of the knee joint were exercised both actively and passively and were continued after discharge. Patients were discharged when they were able to walk with two crutches or any further assistance. Postoperative rehabilitation consisted of six weeks of 50% weight bearing. The follow-up period of the study was two years.

Outcomes

The primary outcome was the difference in accuracy of performed correction between the groups. The accuracy of performed correction was assessed on loaded whole leg radiographs with the hip-knee-ankle (HKA) angle at baseline and six weeks. The HKA angle is defined as the angle between two prolonged lines: one line of the center from

the femur head to the top of the femoral notch and a second line from the center of the ankle to the center of the tibial spines (Figure 1a) [9]. For example, with a pre-operative HKA angle of 3° varus, and a planned correction of 3° valgus, the planned correction was 6°. If postoperative radiographs show an HKA angle of 4° valgus, the performed correction was 7°, which deviates 1° from the planned correction. The accuracy of performed correction was + 1°.

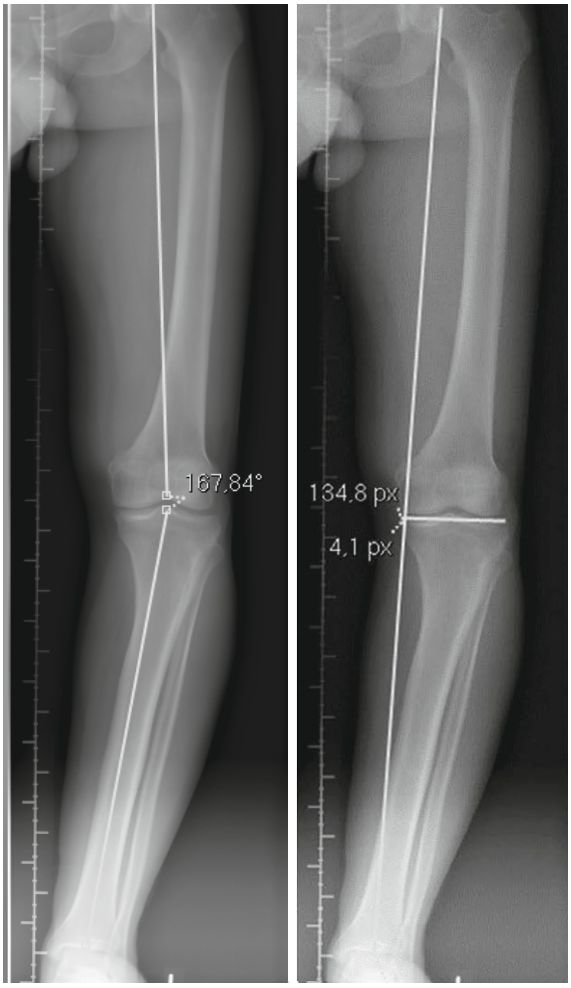


Figure 1a) measurement of hip-knee-ankle angle. **1b)** measurement of position of mechanical axis.

Whole leg radiographs were made according to a pre-specified protocol, for exact anterior-posterior projection of the knee. The standardized whole leg radiographs were made in a standing weight-bearing position[10]. First, a lateral radiograph of the knee was made. In this lateral radiograph, the backside of the femur condyles should exactly

overlap each other. Then, the beaming head was moved perpendicular and the anterior-posterior whole leg radiograph was made[11].

Secondary outcomes were: 1) differences between the groups in HKA-angle and position of mechanical axis at six weeks, 2) differences between the groups in loss of correction between six weeks and two years assessed with HKA-angle and position of mechanical axis, 3) differences in complications between the groups. The position of mechanical axis is defined as the distance of the mechanical axis relative to the medial compartments' medial edge, expressed as a percentage of the total tibial plateau width (Figure 1b) [9]. All measurements were done by the same observer (AW). Thirty randomly selected radiographs were assessed by a second observer (EvE) to assess interrater reliability for HKA angle and position of the mechanical axis.

Complications that were registered during the study period were wound infection, removal of osteosynthesis material, bleeding, numbness of the lower leg, irritation of osteosynthesis, or other complications. Wound infection and removal of osteosynthesis were reported by the treating physician. Patients were asked by questionnaires whether they had experienced bleeding, numbness of the lower leg or irritation of osteosynthesis, or other complications.

For the sample size calculation, the study of Hankemeier et al. was used, in which the accuracy of the open and closed wedge technique was assessed[12]. The target of both techniques was a 3-4° over-correction of the malalignment measured on a whole leg radiograph. The mean deviation of open and closed wedged techniques compared to the preoperative planning was 1.7° (SD ± 1.6°) and 2.6 (SD ± 1.8°) respectively. The sample size of this study is calculated to detect the difference between both groups in deviation of the preoperative planning. Using a power of 80% and α of 0.05 the required sample size is 56 patients per group, for a total of 112 patients. The final total sample size required is 124 patients, to accommodate a 10% potential dropout rate over 1 year.

Statistical analysis

Interrater reliabilities for HKA-angle and position of the mechanical axis were analyzed with intraclass correlation coefficient (ICC) by a two-way random-effect model with absolute agreement. ICC scores were interpreted as poor (ICC <0.5), moderate (ICC 0.5-0.75), good (ICC 0.75-0.9), and excellent (ICC >0.9) reliability[13,14].

Linear mixed models were used to evaluate the difference between the groups in accuracy of performed correction at six weeks. Accuracy of performed correction was the dependent variable and the randomization allocation the independent variable.

Potential confounders (age, body mass index, and sex) were included in the model as fixed effects if they changed the effect estimate by more than 10%. Randomization was stratified for orthopedic surgeon. Therefore orthopedic surgeon was added as a random factor to the model. Differences in HKA angle and position of mechanical axis at six weeks were also analyzed with linear mixed models, as described above.

The difference in loss of correction between open wedge HTO and closed wedge HTO was evaluated by a linear mixed model. Difference in loss of correction between the groups was the dependent variables. Randomization allocation was the independent variable. Again, potential confounders (age, body mass index, and sex) were included in the model if they changed the effect estimate by more than 10%. Orthopedic surgeon was added as random effect.

Logistic regressions were used to analyze differences in complications between the group. Complications were wound infection, removal of osteosynthesis material, bleeding, numbness of the lower leg, irritation of osteosynthesis, or other complications. The complications were the dependent variables. Randomization allocation was the independent variable and orthopedic surgeon was added as a random effect.

Models were checked for linearity, homoscedasticity, and normality of the residuals. Datasets were encrypted until completion of analysis to keep the statistician blinded for randomization outcome. Statistical significance was set at the two-sided 0.05 level.

Results

Patients

Between July 2012 and July 2020, 78 patients were included in this RCT. The follow-up ended in July 2022. Two patients canceled the surgery while they were on the waiting list and were not randomized. Seventy-six patients were randomized, 38 patients were randomized to open wedge HTO, and 38 to closed wedge HTO. In the open wedge HTO group, one patient canceled the surgery and one patient withdraw consent for study participation. In the closed wedge HTO group, the surgery of one patient was cancelled because of severe coronary artery disease. All remaining patients were operated according to the randomized treatment. During the study period, the radiographic protocol was not always applied, due to miscommunication with the radiological department. 17% of the whole leg radiographs were made without a lateral view, and therefore it is uncertain whether these radiographs were made exactly anterior-posterior.

Therefore, we performed an additional sensitivity analysis including only radiographs that were made according to the radiographic protocol. See Figure 2 for a flow chart of the patient selection and Table 1 for the baseline characteristics of the included patients. Although inclusion criteria were verbally checked before inclusion, two patients reported at the baseline questionnaire to have complaints between 1 and 3 months, which was shorter than one of the inclusion criteria (3 months). However, considering the length of the surgical waiting list, none of the included patients had less than 3 months of complaints at the time of surgery.

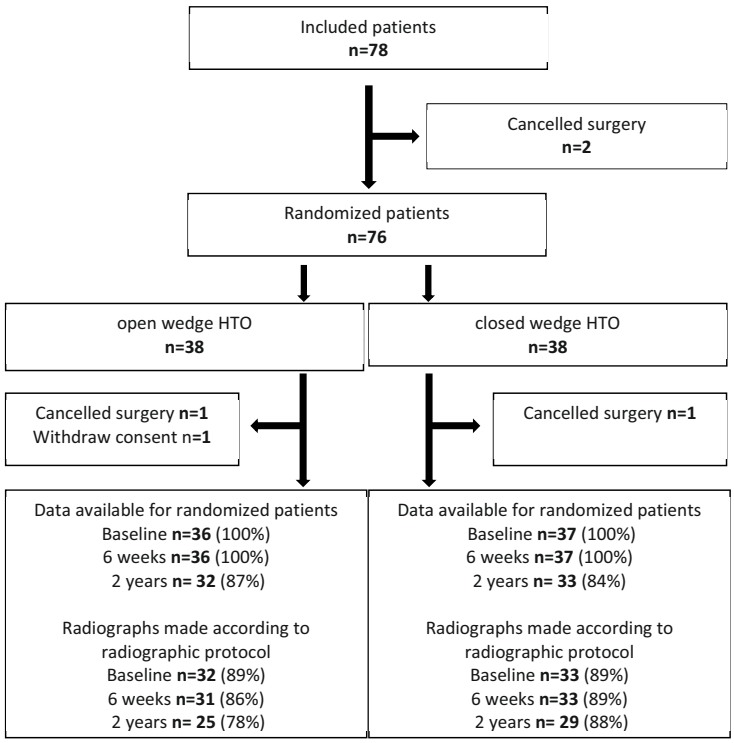


Figure 2. Flow chart of patient selection

Table 1. Baseline characteristics of the study population

	Open wedge HTO (n=38)	Closed wedge HTO (n=38)
Age at inclusion, years	51 (9)	54 (7)
No (%) men	24 (63)	24 (63)
No (%) left knee	15 (40)	19 (50)
Body mass index, kg/m^2	29 (4)	30 (4)
No (%) patients with Kellgren and Lawrence score (%)		
- 1	9 (24)	4 (11)
- 2	9 (24)	16 (42)
- 3	20 (52)	18 (47)
No (%) duration of complaints		
- 1-3 months	1 (3)	1 (3)
- 3-6 months	3 (8)	3 (8)
- 6-12 months	6 (16)	8 (21)
- >12 months	28 (74)	26 (68)
Days with pain in the last month	27 (7)	26 (8)
VAS pain score rest	60 (20)	60 (20)
VAS pain score activity	80 (10)	80 (10)
Hip-knee-ankle angle in degrees*	7 (3)	7 (3)
Position of mechanical axis (%)	18 (16)	20 (14)

Data are presented as mean (standard deviation) unless otherwise indicated; No=number
* Positive values represent varus alignment; negative values represent valgus alignment;

Spaghetti plots of the raw data for the HKA-angle and position of mechanical axis can be found in appendix 1.

Interrater reliability

The ICC was excellent for the HKA-angle (ICC=0.99) and for the position of the mechanical axis (ICC=0.99).

Primary outcome

Compared to pre-operative planning, the open wedge HTO group showed an under-correction (less valgus) of 0.9° (95% confidence interval (CI) -2.0 to 0.1). The closed wedge HTO group showed an over-correction (more valgus) of 1.0° (95% CI -0.1 to 2.0). The difference of 1.9° between the groups was statistically significant (95% CI 0.7 to 3.1) (Table 2).

Secondary outcome

HKA angles at six weeks were significantly different between the groups, with a mean difference of 1.5° (95% IC -2.7 to -0.3). The position of the mechanical axis was also significantly different between the groups, with a mean difference of 7.6% (95% CI 2.5 to 12.6) (Table 2, Figure 3). Loss of correction between six weeks and two years was not statistically significant between the groups assessed with HKA and position of mechanical axis (Table 3, Figure 3).

The results of the sensitivity analysis did not show different results than the primary analysis (Table 2 and 3).

No significant differences between both groups were found for occurrence of complications (wound infection, removal of osteosynthesis, bleeding, numbness of the lower leg, plate irritation, and other complications) (Table 4). Other complications that were reported by the patients were for instance: pain in hip, back, knee or lower leg, swelling of the knee or instability of the knee.

Table 2. Correction at six weeks

	Open wedge HTO	**Closed wedge HTO**	**Mean difference[1]**
Primary analysis			
Accuracy of performed correction in degrees	-0.9 (-2.0 to 0.1)	1.0 (-0.1 to 2.0)	1.9 (0.7 to 3.1)*
Hip-knee-ankle angle at six weeks in degrees[3]	-2.8 (-3.9 to -1.6)	-4.3 (-5.4 to -3.2)	-1.5 (-2.7 to -0.3)*
Position of mechanical axis at six weeks in percentages	60.9 (56.2 to 65.5)	68.4 (63.9 to 72.9)	7.6 (2.5 to 12.6)*
Sensitivity analysis			
Accuracy of performed correction in degrees[2]	-0.9 (-2.1 to 0.2)	0.9 (-0.2 to 2.1)	1.9 (0.6 to 3.1)*
Hip-knee-ankle angle at six weeks in degrees[3]	-2.9 (-4.1to -1.7)	-4.2 (-5.3 to -3.0)	-1.3 (-2.6 to -0.004)*
Position of mechanical axis at six weeks in percentages	60.9 (56.2 to 65.5)	68.4 (63.9 to 72.9)	7.5 (1.0 to 12.0)*

Data are mean (95% confidence interval); * Significant different between the groups with p<0.05
[1]Adjusted for potential confounders (gender, age, body mass index, orthopedic surgeon)
[2]Positive values indicate over-correction; negative values indicate under-correction
[3]Positive values represent varus alignment; negative values represent valgus alignment

Table 3. Correction loss between 6 weeks and 2 years sensitivity analysis

	Open wedge HTO	Closed wedge HTO	Mean difference[1]
Primary analysis			
Loss of correction in HKA in degrees	0.5 (-0.4 to 1.4)	1.0 (0.2 to 1.9)	-0.5 (-0.4 to 1.5)
Loss of correction in PMA in percentages	-1.9 (-5.5 to 1.7)	-4.5 (8.1 to -0.8)	-2.6 (-6.5 to 1.4)
Sensitivity analysis			
Loss of correction in HKA in degrees	0.8 (-0.3 to 1.8)	1.0 (0.04 to 2.0)	-0.2 (-1.0 to 1.4)
Loss of correction in PMA in percentages	-2.4 (-6.7 to 2.0)	-4.3 (-8.5 to -0.2)	2.0 (-6.7 –to 2.8)

Data are means (95% confidence interval)
HKA: hip-knee-ankle angle; PMA: position of mechanical axis.
[1] Adjusted for possible confounders (gender, age, body mass index, orthopedic surgeon);

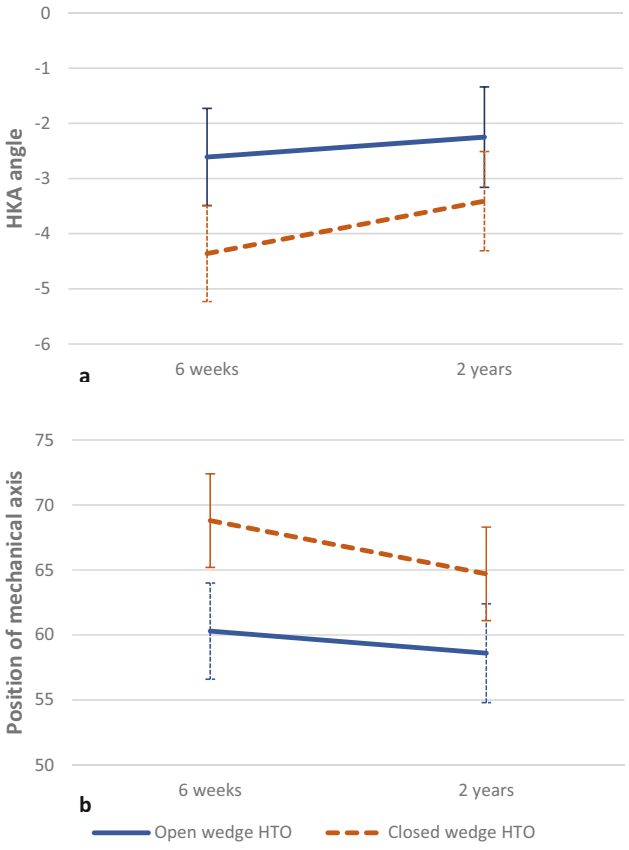


Figure 3a) Loss of correction in HKA (hip-knee-ankle) angle. **3b)** Loss of correction in position of mechanical axis

Discussion

This RCT compared the accuracy of performed correction between open and closed wedge HTO. The results showed that performed corrections were significantly different between open and closed wedge HTO, with a non-significant under-correction for open wedge HTO and non-significant over-correction for closed wedge HTO. Loss of correction between six weeks and two years was not significantly different between the groups. Also, no significant differences were found in complications between the two techniques.

It has been hypothesized that open wedge HTO can achieve more accurate correction due to the possibility of fine adjustment. This hypothesis was strengthened by the studies of Hankemeier et al. (2010) and the RCT of Luites et al. (2009), who reported more accurate correction with open wedge HTO [12,15]. On the contrary, there also are multiple RCTs which report no differences in post-operative HKA angle or accuracy of correction [16,17]. Two meta-analyses were performed that compared open and closed wedge HTO [6,18]. These studies did not find any significant differences in accuracy of correction between the two techniques [6,18]. Based on our study results and the available literature, it seems that open and closed wedge HTO can both achieve accurate correction. Therefore, both techniques seem to be equally good in achieving the planned correction.

Our study showed non-significant under-correction for open wedge HTO (0.9°; 95% CI -2.0 to 0.1). and non-significant over-correction for closed wedge HTO (1.0°; 95% CI -0.1 to 2.0). Although the deviations from the planned correction were relatively small, they may affect long-term results of HTO. It is currently unknown what the long-term effects of HTO over- and under-correction are. It seems likely that under-correction could result in insufficient unloading of the medial compartment, whereas over-correction may lead to overloading of the lateral compartment. Also, no cut-off values exist for which deviation from the planned correction is still acceptable. For future studies, it would be interesting to focus on long-term effects of HTO over- and under-correction, and to see to what extent this affects clinically important outcomes.

In our study, we did not find significant differences in loss of correction between the groups over two years, which indicates that both techniques can accurately stabilize an HTO. However, we did not perform a power calculation for this secondary outcome measure and therefore possible differences may not have been detected due to a lack of power. Also, we did not find any significant differences in complications between open and closed wedge HTO. This could also be caused by a lack of power but it does seem to indicate that both techniques are equally safe. The study of Hankemeier et al. (2010) also reported no significant differences in the number of complications after open and closed

wedge HTO[12]. However, this study reported no data on the frequency of osteosynthesis removal. The study of Van Egmond et al. (2016) did not find any significant differences in the frequency of hardware removal, whereas the study of Brouwer et al. (2006) did report more frequently osteosynthesis removal after open wedge HTO (60%) compared to closed wedge HTO (23%) [19]. This was also found in the study of Duivenvoorden et al. (2017)[20]. Wang et al. (2018) concluded in their meta-analysis that there is no significant difference in complications between open and closed wedge HTO. Overall, it seems that complication rates after open and closed wedge HTO are similar.

Previous studies have shown that a combination of knee flexion and simultaneous leg rotation during the whole leg radiograph influences coronal alignment measurements [21]. This knee position is frequently seen in patients with total knee arthroplasty during the early postoperative period [21]. As knee flexion and leg rotation may also occur in the early postoperative phase after HTO, a strict radiological protocol was used to minimize the effects of flexion and rotation on our radiological measurements. Unfortunately, the protocol was not always followed. 17% of the whole leg radiographs were made without a lateral radiograph and therefore rotation of the knee could be present. We performed a sensitivity analysis to see whether this protocol violation influenced our results. The sensitivity analysis did not lead to different results compared to our primary analysis.

A strong feature of this study is that it is a RCT. This study design limits risk of bias and results in high-quality evidence. However, the study has some limitations. Patients, orthopedic surgeons, and radiographic assessors were not blinded for the randomization outcome. Patients were able to see which technique was used based on position of the wound/scar and on their radiographs. However, it seems unlikely that this knowledge biased the results of this RCT as our primary outcome was based on measurements of radiographs. The assessor of the radiographs was not blinded, as it was impossible to remove osteosynthesis material from the radiographs. However, the assessor was independent without a preference for one of the techniques. Also, statistical analysis was done with encrypted data to keep the risk of bias reduced as much as possible.

To check whether the performed correction was accurate, a metal rod was used to mimic the mechanical axis.

Our sample size calculation indicated that we needed to include 124 patients to find significant differences between open and closed HTO. However, due to slow inclusion rate and the worldwide COVID-19 pandemic, we were forced to stop the inclusion at 76 included patients. Nevertheless, we did find significant differences for our primary outcome measure.

Conclusion

In this study, the accuracy of performed correction between open and closed wedge HTO was compared. Although the accuracy from performed correction was significantly different between the groups, both techniques seem to achieve an accurate correction. It is currently unknown what the long-term effects of over- or under-correction are on the progression of knee OA. Therefore, based on correction accuracy in our study, we conclude that both techniques can be used to perform HTO for varus malalignment.

References

1 Robertsson, O., Lidgren, L., Sundberg, M. & W-Dahl, A. *The Swedish Knee Arthroplasty Register - Annual report 2020*. (2020).

2 Sharma, L. *et al.* The role of knee alignment in disease progression and functional decline in knee osteoarthritis. *Jama* **286**, 188-195 (2001).

3 Kanakamedala, A. C. *et al.* High Tibial Osteotomies for the Treatment of Osteoarthritis of the Knee. *JBJS Rev* **10** (2022).

4 Amendola, A. & Bonasia, D. E. Results of high tibial osteotomy: review of the literature. *Int Orthop* **34**, 155-160 (2010).

5 Ollivier, B., Berger, P., Depuydt, C. & Vandenneucker, H. Good long-term survival and patient-reported outcomes after high tibial osteotomy for medial compartment osteoarthritis. *Knee Surg Sports Traumatol Arthrosc* **29**, 3569-3584 (2021).

6 Wang, Z., Zeng, Y., She, W., Luo, X. & Cai, L. Is opening-wedge high tibial osteotomy superior to closing-wedge high tibial osteotomy in treatment of unicompartmental osteoarthritis? A meta-analysis of randomized controlled trials. *Int J Surg* **60**, 153-163, doi:10.1016/j.ijsu.2018.10.045 (2018).

7 Sun, H., Zhou, L., Li, F. & Duan, J. Comparison between Closing-Wedge and Opening-Wedge High Tibial Osteotomy in Patients with Medial Knee Osteoarthritis: A Systematic Review and Meta-analysis. *J Knee Surg* **30**, 158-165, doi:10.1055/s-0036-1584189 (2017).

8 Brouwer, R. W. *et al.* Osteotomy for treating knee osteoarthritis. *Cochrane Database Syst Rev* **2014**, CD004019 (2014).

9 Higuchi, T. *et al.* Comparison of radiological features of high tibial osteotomy and tibial condylar valgus osteotomy. *BMC Musculoskeletal Disorders* **20**, 409, doi:10.1186/s12891-019-2764-0 (2019).

10 Brouwer, R. W., Jakma, T. S., Bierma-Zeinstra, S. M., Ginai, A. Z. & Verhaar, J. A. The whole leg radiograph: standing versus supine for determining axial alignment. *Acta Orthop Scand* **74**, 565-568 (2003).

11 Brouwer, R. W., Jakma, T. S., Brouwer, K. H. & Verhaar, J. A. Pitfalls in determining knee alignment: a radiographic cadaver study. *J Knee Surg* **20**, 210-215 (2007).

12 Hankemeier, S. *et al.* Accuracy of high tibial osteotomy: comparison between open- and closed-wedge technique. *Knee Surg Sports Traumatol Arthrosc* **18**, 1328-1333 (2010).

13 Koo, T. K. & Li, M. Y. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med* **15**, 155-163 (2016).

14 Portney, L. G. & Watkins, M. P. *Foundations of clinical research: applications to practice*.

(Pearson/Prentice Hall Upper Saddle River, NJ, 2009).

15 Luites, J. W., Brinkman, J. M., Wymenga, A. B. & van Heerwaarden, R. J. Fixation stability of opening- versus closing-wedge high tibial osteotomy: a randomised clinical trial using radiostereometry. *J Bone Joint Surg Br* **91**, 1459-1465 (2009).

16 Nerhus, T. K. *et al.* Radiological outcomes in a randomized trial comparing opening wedge and closing wedge techniques of high tibial osteotomy. *Knee Surg Sports Traumatol Arthrosc* **25**, 910-917 (2017).

17 Kim, J. I. *et al.* Lower Limb Length Discrepancy After High Tibial Osteotomy: Prospective Randomized Controlled Trial of Lateral Closing Versus Medial Opening Wedge Osteotomy. *The American Journal of Sports Medicine* **44**, 3095-3102, doi:10.1177/0363546516659284 (2016).

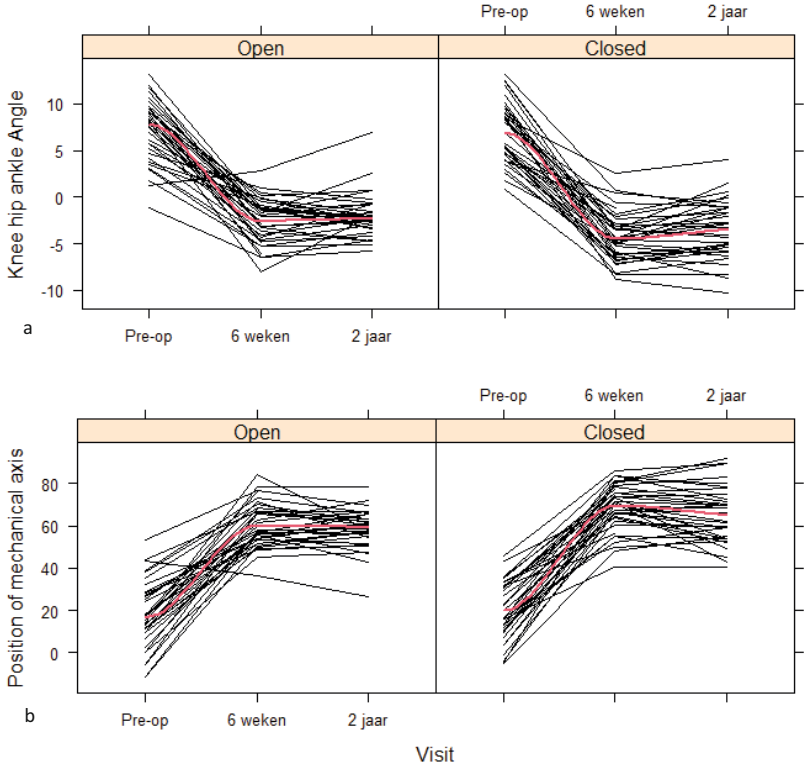
18 Cheng, X., Liu, F., Xiong, F., Huang, Y. & Paulus, A. C. Radiographic changes and clinical outcomes after open and closed wedge high tibial osteotomy: a systematic review and meta-analysis. *J Orthop Surg Res* **14**, 179 (2019).

19 Brouwer, R. W., Bierma-Zeinstra, S. M., van Raaij, T. M. & Verhaar, J. A. Osteotomy for medial compartment arthritis of the knee using a closing wedge or an opening wedge controlled by a Puddu plate. A one-year randomised, controlled study. *J Bone Joint Surg Br* **88**, 1454-1459 (2006).

20 Duivenvoorden, T. *et al.* Adverse events and survival after closing- and opening-wedge high tibial osteotomy: a comparative study of 412 patients. *Knee Surg Sports Traumatol Arthrosc* **25**, 895-901 (2017).

21 Yoo, H. J. *et al*. Pitfalls in assessing limb alignment affected by rotation and flexion of the knee after total knee arthroplasty: Analysis using sagittal and coronal whole-body EOS radiography. *The Knee* **27**, 1551-1559, doi:https://doi.org/10.1016/j.knee.2020.08.008 (2020).

Appendix



Appendix 1: a) Raw data for knee hip-knee-ankle angle for open wedge HTO and closed wedge HTO; red line indicates average hip-knee-ankle angle.
b) Raw data for position of mechanical axis for open wedge and closed wedge HTO; red line indicates average position of mechanical axis.

Chapter 3

Talocrural arthrodesis increases osteoarthritis severity in adjacent joints: a midterm computed tomography follow-up study

A. Willems, M. Minnaard, E.H.G. Oei, S.M.A. Bierma-Zeinstra, D.E. Meuffels
Submitted

Abstract

Background: After talocrural arthrodesis, adjacent joints (subtalar, talonavicular and calcaneocuboid) are often affected by osteoarthritis (OA). It is unclear if OA is pre-existing to talocrural arthrodesis, or whether it develops after talocrural arthrodesis. This retrospective study is unique because it is the first study with pre-operative and follow-up computed tomography (CT).The aim of this study is to investigate whether OA develops in adjacent joints after talocrural arthrodesis or if OA is already pre-existing. In addition, associations between degree of OA and patient reported outcomes are investigated.
Methods: Patients were selected from electronic files and adjacent joint OA was assessed on preoperative CT and bilateral follow-up CT. Patient reported outcomes were collected.
Results: Twenty-three patients were included with an average follow-up time of 7 years (SD=2). In participants without pre-existing OA, OA significantly progressed in all adjacent joints. In participants with pre-existing OA, OA progressed in the subtalar joint. Patient reported outcomes were not correlated to OA.
Conclusions: OA in the adjacent joints progresses after talocrural arthrodesis, especially in participants without pre-existing OA. The severity of OA is not related to patient reported outcomes. Therefore, the clinical impact of the progression of OA seems to be limited.
Level of Evidence: Level III, retrospective

Introduction

A talocrural arthrodesis can be a life-changing operative intervention for patients with end-stage talocrural osteoarthritis (OA). After talocrural arthrodesis, pain scores are significantly reduced resulting in a better quality of life [1] . However, mid- and long-term follow-up studies showed that OA is present in adjacent joints (subtalar, talonavicular and calcaneocuboid joints) after talocrural arthrodesis [2]. It has been hypothesized that these arthritic changes develop after talocrural arthrodesis as a consequence of increased use and higher forces in the joints [3,4]. In contrast, it has been suggested that arthritic changes are pre-existing to talocrural arthrodesis [5]. Although adjacent joint OA after arthrodesis is a widely studied subject [6-11], a paucity of studies evaluate the pre-operative prevalence of adjacent joint OA[2]. It is therefore unclear whether adjacent joint OA is present before talocrural arthrodesis or if it develops after talocrural arthrodesis [2].

Also, so far all studies describing adjacent joint OA after talocrural arthrodesis have used radiographs to assess OA[2,6-11]. The assessment of OA in the subtalar and talonavicular joints from radiographs with Kellgren-Lawrence score showed poor reliability, as radiographs lack bony details for reliable OA assessment [12,13]. Alternatively, computed tomography (CT) provides cross-sectional images, from which all parts of the joints can be assessed in detail and are more accurate [14].

This study presents a cohort of patients that underwent preoperative and postoperative CT. It is unique because both preoperative and postoperative CT is available, thus development or progression of OA in adjacent joints can be precisely determined. We aim to assess whether adjacent joint OA is present before talocrural arthrodesis, or if it develops postoperatively. Furthermore, we will also correlate OA in adjacent joints to length of follow-up, patient reported outcome measures, and measures of patients' satisfaction.

Methods

Study design and participants

This retrospective cohort presents data of patients who underwent talocrural arthrodesis at the Erasmus MC University Medical Centre, Rotterdam, the Netherlands between January 2008 and June 2016. Patients were indicated for an isolated talocrural arthrodesis if they experienced symptomatic talocrural OA, sometimes accompanied with postural deviations of the ankle, and without complaints of the adjacent joints and postural

deviations in the foot. An electronic search in the hospital files was performed based on operative codes, to select all patients who underwent a talocrural arthrodesis. After the electronic search, patients were screened for eligibility based on predefined in- and exclusion criteria (Figure 1). Eligible patients were approached to participate in the study. Patients who were willing to participate gave written informed consent and were invited to the outpatient clinic for follow-up examination. Follow-up examination included bilateral CT and completion of patient reported outcome measures.

Inclusion criteria	Exclusion criteria
• Primary talocrural arthrodesis performed between 1-1-2008 until 30-06-2016 between at the department of Orthopaedics of the Erasmus MC • Minimal follow-up period of 3 years • Preoperative CT of the operated ankle is available	• Double or triple arthrodesis • Amputation of the ankle after arthrodesis • Patient deceased • Patients < 18 years at time of ankle arthrodesis • CT > 1 year before surgery.

Figure 1. Study inclusion and exclusion criteria

Outcome measures
Baseline characteristics

Baseline characteristics were extracted from the electronic patient file (age, body mass index (BMI), gender, operated side, reason for talocrural arthrodesis, technique used for talocrural arthrodesis, and date of talocrural arthrodesis). At follow-up examination participants were asked whether they experienced problems with the contralateral ankle.

Grading OA of tarsal joints with CT OA scale

The degree of OA in the adjacent joints was assessed on the preoperative and on bilateral follow-up CT, where the non-affected ankle served as a control. The degree of OA in the adjacent joints was assessed with a modified assessment tool based on the Kellgren-Lawrence OA scoring (0-4) and CT ankle OA atlas [14,15]. This modified tool, the CT OA scale, contains four features which are associated with OA: subchondral sclerosis, cysts, joint space narrowing and osteophytes. Each feature was scored separately on a scale from 0 to 3, where 0 is absence of the feature and 3 the worst severity of the feature. Figure 2 shows the grading of the CT OA scale. After all features were scored, a total score per joint was calculated by totalling all scores of the individual features. The total score ranges between 0 representing no OA, and 12 representing worst OA. Joints were assessed based on multiplanar reconstructions in three planes (sagittal, coronal and axial), with slice thickness of 1 mm. All CT's were scored by one observer (AW). A random sample of 12 CT's was scored by a second observer (DM) to assess interrater reliability for each feature and for the overall score.

	0	1	2	3
1. Subchondral sclerosis	No subchondral sclerosis	Mild subchondral sclerosis	Sclerosis	Significant subchondral sclerosis
2. Cysts	No cysts	No obvious subchondral cysts	Obvious subchondral cyst formation	Significant cyst formation
3. Joint space narrowing	Joint space integrity fully intact	Possible mild joint space narrowing	Near joint space narrowing	Absence of joint space
4. Osteophytes	No spurring	Mild osteophyte formation / lipping (spurring) present	Obvious osteophyte formation, multiple osteophyte formation	Large osteophytes

Figure 2. CT OA scale: scoring tool for OA in ankle and foot joints from CT

After scoring CT's with CT OA scale, observers were asked to indicate whether, in their opinion, OA was present in the joint or not (yes or no). These outcomes were used to set a cut-off value for the OA ankle scale to discriminate between joints with pre-existing OA and joints without pre-existing OA.

Patient reported outcome measures

Patient reported outcome measures were assessed with three questionnaires. The '36-Item Short Form Health Survey' (SF-36) about quality of life, containing eight subdomains (physical functioning, limitations due to physical health, limitations due to emotional problems, energy/fatigue, emotional well-being, social functioning, pain and general health) [16]. The Foot and Ankle Outcome Score (FAOS) assesses ankle function and pain [17]. The American Orthopaedic Foot and Ankle Society(AOFAS) Ankle-Hindfoot scale assesses pain and impairment of the ankle [18]. The AOFAS Ankle-Hindfoot scale has a clinical reported part also, which was completed by the Orthopaedic surgeon in residence. The scores of all three patient reported outcome measures range between 0 and 100, where a higher score indicates better quality of life (SF-36) or better ankle function (FAOS and AOFAS Ankle-Hindfoot score).

Satisfaction

The degree of satisfaction was scored by numeric rating scale (NRS), where 0 indicated 'very dissatisfied' and 10 'very satisfied'. In addition, participants were asked if they would choose talocrural arthrodesis again.

Statistical analysis

Data were analysed using IBM SPSS Statistics 25. Statistical significance was set at $p<.05$. Data were checked for normality by performing Shapiro-Wilk tests and visual inspection of the Q-Q plots.

Patients without preoperative CT were excluded from the study. The baseline characteristics of these patients were compared to the baseline characteristics of the included participants with unpaired t-test and chi-square test.

Interrater reliability for the CT OA scale were analysed by a two-way random-effect model with absolute agreement. Interrater reliability was assessed on a random subset of 12 CT's. Interrater reliability scores were interpreted according to the Koch-Landis method, in which kappa (κ) scores can be interpreted as indicating slight agreement (k=.01-.2), fair agreement (k=.21-.40), moderate agreement (k=.41-.60), substantial agreement (k=.61-.80) and excellent agreement (k=.81-1.00) [19].

To assess differences in OA between preoperative and follow-up CT's for each adjacent joint, three paired t-tests were performed per joint with Bonferroni-Holm adjustment to correct for multiple testing for each adjacent joint.

Per adjacent joint, participants were grouped based on pre-existing OA. The cut-off value for pre-existing arthritis was set by ROC analysis and Youden index. In this analysis, the optimal cut-off value for OA ankle scale is set by relating OA ankle scores to the observers' assessment of OA being present or not present in the joint. The ROC analysis calculates sensitivity and specificity scores for all possible cut-off values [20]. Youden index is calculated by 'sensitivity+specificity-1'. The highest Youden index indicates the optimal cut-off value, i.e. optimal balance between sensitivity and specificity [21].

To investigate increase in OA per group, paired t-tests were performed in case of normal distribution, in case of non-normal distribution Wilcoxon signed rank test were done.

To assess whether OA development is associated with time after talocrural arthrodesis, multiple regression analysis was performed for each joint, which were adjusted for age. Also, multiple regression analysis with age adjustment was used to investigate associations between degree of OA at follow-up and patient reported outcomes (SF-36, FAOS, AOFAS Ankle-Hindfoot score and satisfaction).

The study was approved by the local Medical Ethical Committee (MEC-2018-153). No funding was received and the authors have no conflicts of interest to declare.

Results

General characteristics of study population

The electronic search for eligible participants resulted in 83 potentially eligible participants. After screening, 27 eligible patients remained. These patients were approached to participate in the study. Eventually, 23 patients agreed to participate and were included in the study. See Figure 3 for a flow-chart with exclusion reasons for the other patients.

14 patients were excluded because of missing preoperative CT. Baseline characteristics were compared between included patients and patients with missing preoperative CT to investigate possible selection bias . The characteristics age, BMI, operated side, type of OA and operative technique were not significantly different between patients with and without CT. Time since surgery and gender were significantly different between the groups. Table 1 shows the baseline characteristics of the included participants and characteristics of patients that were excluded due to missing pre-operative CT.

9 participants reported problems of the contralateral ankle at follow-up. 3 participants had a talocrural arthrodesis, 3 participants reported persistent pain of the ankle after a fracture or trauma, 1 participant had Achilles tendon lengthening, 1 participant had OA in the ankle joint, and 1 participant reported mild ankle complaints.

The three participants with a contralateral arthrodesis were excluded from the analysis. Furthermore, 1 participant had an ipsilateral subtalar arthrodesis during follow-up period and 1 pre-operative CT did not include the calcaneocuboid joint. Therefore, degree of OA could not be measured on these CT's.

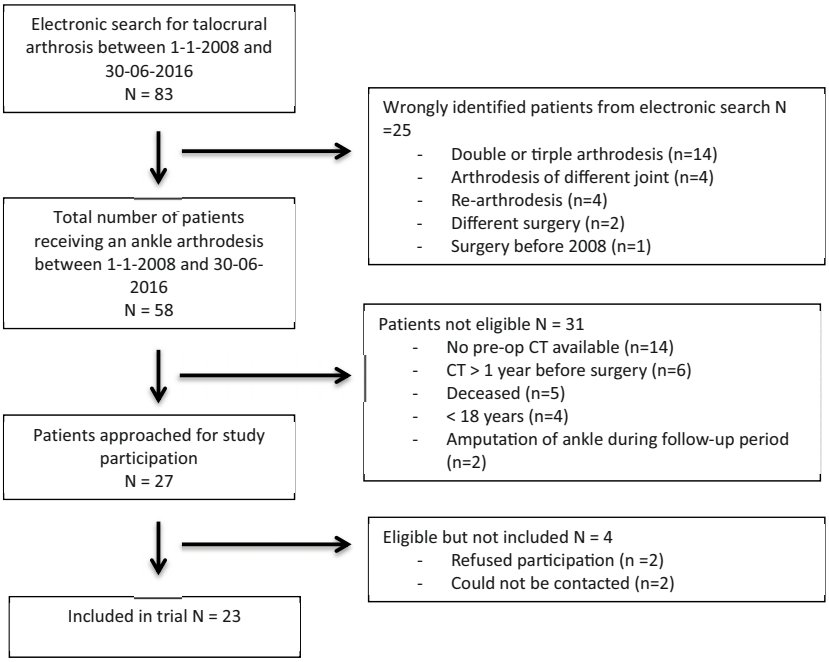


Figure 3. Flowchart of patient selection

Table 1. Characteristics of included patients and patients excluded due to missing preoperative CT

	Included patients (n=23)	**Excluded patients due to missing CT (n=14)**	**P-value**
Average follow-up period in years (SD)	7 (2)	9 (2)	0.005*
Average age at arthrodesis in years (SD)	52 (15)	52 (17)	1.0
Average BMI (SD)	28 (4)	28 (6)	0.9
Men	10 (44)	11 (79)	0.03*
Talocrural arthrodesis at right side	13 (57)	7 (50)	0.7
Type of osteoarthritis			0.4
Primary	3 (13)	4 (31)	
Secondary	20 (87)	10 (69)	
Operative technique			0.8
Open arthrodesis	4 (17)	3 (21)	
Arthroscopic arthrodesis	19 (83)	11 (79)	

Data are absolute numbers (%) unless otherwise indicated; SD: Standard deviation; CT: computed tomography; BMI: body mass index; * Significantly different with $p<0.05$

Interrater reliability for degree of OA

Interrater reliability was excellent for all features separately and for the overall score (sclerosis k=0.86 (95% confidence interval (CI) 0.6-0.9); cysts k=0.87 (95%CI 0.7-0.9); joint space narrowing k=0.82 (95% CI 0.6-0.9); osteophytes k=0.94 (95% CI 0.9-1.0); overall score k=0.94 (95% CI 0.9-1.0))

Degree of OA in adjacent joints

Overall, the CT OA scale scores were significantly higher for the adjacent joints of the operated ankle at follow-up, compared to preoperative situation and compared to these joints in the control feet. The CT OA scale scores between the preoperative CT and follow-up CT in the control group were not significantly different for any of the adjacent joints, see Table 2.

Table 2. CT OA scale score in adjacent joints before and after ankle arthrodesis

	Preoperative operated	**Follow-up operated**	**P-value**
Subtalar joint (n= 22)[A]	4.7 (3.0)	7.2 (2.4)	<0.001*
Talonavicular joint (n= 23)	3.1 (1.9)	4.8 (2.2)	0.003*
Calcaneocuboid joint (n= 22)[B]	2.1 (1.6)	4.1 (1.7)	<0.001*
	Preoperative operated	Follow-up control	
Subtalar joint (n= 20)[C]	4.4 (3.0)	3.1 (2.5)	0.3
Talonavicular joint (n= 20)[C]	2.9 (1.7)	2.5 (2.0)	1.0
Calcaneocuboid joint (n= 19)[B,C]	2.1 (1.7)	1.7 (2.1)	1.0
	Follow-up operated	Follow-up control	
Subtalar joint (n= 19)[A,C]	7.0 (2.5)	3.2 (2.5)	<0.001*
Talonavicular joint (n= 20)[C]	4.9 (2.3)	2.5 (2.0)	0.006*
Calcaneocuboid joint (n= 20)[C]	4.2 (1.6)	1.6 (2.0)	<0.001*

Data are mean (standard deviation); * Significantly different with p<.05; A: Data of one patient missing due to ipsilateral subtalar arthrodesis during follow-up. B: Data of one patient missing due to missing calcaneocuboid joint on pre-operative CT; C: Data of three patients missing due to arthrodesis of the control ankle;

Pre-existing OA

Optimal cut-off value for OA ankle scale to discriminate between joints with pre-existing OA and without pre-existing OA was 3.5. For this cut-off value, the sensitivity and specificity were 0.9, the Youden index was 0.8.

Table 3 shows the changes in OA in adjacent joints for joints with pre-existing OA and without pre-existing OA. Overall, progression of OA was found in all adjacent joints without pre-existing OA. In adjacent joints with pre-existing OA progression was only found in the subtalar joint. At follow-up, OA was present in 22 participants (96%) in the subtalar joint, in 15 participants (65%) in the talonavicular joint, and in 14 participants (61%) in the calcaneocuboid joint.

Table 3. Changes in CT OA scale score for adjacent joints with pre-existing OA and without pre-existing OA

	Preoperative	**Follow-up**	**p-value**
Subtalar joint			
- Pre-existing OA (n=14)	5 (4-8)	8 (6-9)	0.01*
- No pre-existing OA (n=9)	2 (1-3)	5 (5-8)	0.01*
Talonavicular joint			
- Pre-existing OA (n=9)	5 (4-6)	5 (4-7)	0.6
- No pre-existing OA (n=14)	2 (1-2)	4 (3-6)	0.002*
Calcaneocuboid joint			
- Pre-existing OA (n=3)	5 (4-6)	6 (5-7)	0.2
- No pre-existing OA (n=19)	2 (1-3)	4 (2-5)	0.001*

Data are median (interquartile range); * Significantly different with p<.05

Patient reported outcomes and correlations with OA

No significant associations were found between length of follow-up and difference in degree of OA. The SF-36 scores at follow-up are presented in Table 4. For the SF-36, the only domain that was correlated to degree of OA was limitations due to emotional problems. Age-adjusted regression analysis showed significant negative correlations between limitation due to emotional problems and degree of OA at follow-up for the subtalar (r= -0.6, p=0.001) and talonavicular joints (r= -0.6, p=0.001).

The average FAOS at follow-up was 53 (SD=20). The average AOFAS Ankle-hindfoot score at follow-up was 58 (SD=24). No significant associations were found between scores of the FAOS or AOFAS Ankle-Hindfoot score and the degree of OA of the adjacent joints at follow-up, corrected for age.

Table 4. SF-36 scores at follow-up

	Mean (SD)
Physical functioning	54.8 (26.0)
Role limitations due to physical health	56.5 (40.7)
Role limitations due to emotional problems	87.0 (34.4)
Energy/fatigue	65.4 (17.6)
Emotional well-being	80.3 (15.7)
Social functioning	78.3 (22.4)
Pain	56.3 (29.0)
General health	56.3 (23.3)
Health change	45.7 (22.2)

SD: Standard deviation

Patient satisfaction

The median NRS satisfaction score was 8 (interquartile range 7-9). Most participants indicated that they would probably (17,4%) or definitely (69,6%) have the surgery again if they would be asked to choose again. No significant associations were found between satisfaction rate and degree of OA for any of the adjacent joints at follow-up.

Discussion

The overall results of this study showed that the talonavicular, calcaneocuboid, and subtalar joints all showed progression of OA after talocrural arthrodesis. . Degree of OA were not significantly different between controls and pre-operative OA scores. Based on these results, it seems that OA is a result of talocrural arthrodesis.

Coester et al. (2001) compared the degree of OA in adjacent joints between the talocrural side and contralateral side after 22 years of follow-up. In line with the results found in our study, they reported that OA scores were higher at the talocrural side compared to the contralateral side. However, this study had no pre-operative imaging[7].

The studies of Hendrickx et al. (2011), Zwipp et al (2020), Gaedke et al. (2018), Jones et al. (2018) investigated OA before and after talocrural arthrodesis based on radiographs. The cohorts of these studies were very similar to our cohort with average age ranging from 47 to 61 years, and lengths of follow-up between 5 and 10 years and traumatic OA as primary cause for talocrural arthrodesis. Hendrick et al. (2011) reported mild increase in OA in

adjacent joints after talocrural arthrodesis[8]. Zwipp et al. (2010) reported development of OA in 17% of subtalar joints and 11% of talonavicular joints.

Gaedke et al. (2018) reported low pre-existing degrees of OA, which increased after talocrural arthrodesis[9]. This is in accordance with the results found in the present study, which showed that in participants without pre-existing OA, OA progresses significantly after talocrural arthrodesis.

The study of Jones et al. (2018) reported high rates of pre-existing OA , which remained relatively stable during follow-up for the talonavicular joint. Jones et al. (2018) reported that 85% of the patients had no change in talonavicular OA during follow-up[10]. Our study also showed that the degree of OA in participants with pre-existing OA remained relatively stable for the talonavicular and calcaneocuboid joints.

For the subtalar joint, Jones et al. (2018) reported no change in OA in 69% of the patients, and thus an increase in OA in 31% of the patients[10]. Zwipp et al. (2010) also showed progression of preexisting subtalar OA in 30% of the patients. Our study showed progression of OA in participants with pre-existing subtalar joint OA. It therefore seems that preexisting OA in the subtalar joint becomes worse after talocrural arthrodesis.

Overall, the subtalar joint is most affected by OA as 96% of patients have OA in the subtalar joint at follow-up, compared to 65% and 61% for the talonavicular and calcaneocuboid joints.

In this study, patient reported outcomes were measured at follow-up with SF-36, FAOS and AOFAS Ankle-Hindfoot score. For most outcomes there was no association with OA. Other studies could neither find any correlations between SF-36, AOFAS Ankle-Hindfoot score or pain with degenerative changes in adjacent joints [2,7-11]. Our results showed that participants reported high satisfaction rates after talocrural arthrodesis, which were also reported in other studies [7,8,22-24]. Based on the lack of correlations, it seems that increase in radiologically assessed OA does not have a direct impact on patient reported outcomes and satisfaction, and that therefore clinical impact is limited.

The present study has weaknesses. The rate of participants with pre-existing OA was low in our cohort. In absolute numbers, 9 participants had pre-existing OA in the talonavicular joint and 3 in the calcaneocuboid joint. Increase of OA in the adjacent joints may not have been detected due to low statistical power. Low statistical power may also explain why the statistically significant increase in OA was not associated with length of follow-up. Furthermore, selection bias may have occurred due to the relatively large group of

patients that were excluded due to missing preoperative CT, but seems to be limited as baseline characteristics between the groups were comparable. There were some missing data in our cohort. One participant received a subtalar arthrodesis during the follow-up and therefore the degree of OA could not be assessed. This probably led to a slight underestimation of the degree of OA at follow-up in the subtalar joint of the operated ankle. Three control feet were excluded because they had undergone talocrural arthrodesis and could therefore not serve as controls. This may have resulted in an underestimation of degree of OA in the control group. In addition, this is a retrospective cohort without a power calculation. Therefore, non-significant findings and lack of significant associations might be the result of low power. However, despite the above limitations, this is the first study with preoperative CT and postoperative CT which is a strong feature of this study. To draw more firm conclusions, future studies should include higher number of patients. Also for future studies weight bearing CT should be considered as joint space may decrease at weight-bearing, which is missed with standard CT [25].

Conclusion

The present study showed that progression of OA in the adjacent tarsal joints is a consequence of talocrural arthrodesis. Especially adjacent joints without pre-existing OA develop OA after talocrural arthrodesis. The subtalar joint is the most affected by OA with high pre-operative OA scores and progression of OA after talocrural arthrodesis. Progression of OA in adjacent joints does not seem to affect patient reported outcome measures or satisfaction.

References

1. Haddad SL, Coetzee JC, Estok R, Fahrbach K, Banel D, Nalysnyk L. Intermediate and Long-Term Outcomes of Total Ankle Arthroplasty and Ankle Arthrodesis: A Systematic Review of the Literature. *JBJS*. 2007;89(9)
2. Ling JS, Smyth NA, Fraser EJ, et al. Investigating the relationship between ankle arthrodesis and adjacent-joint arthritis in the hindfoot: a systematic review. *J Bone Joint Surg Am*. Mar 18 2015;97(6):513-20.
3. Schuh R, Hofstaetter JG, Hofstaetter SG, Adams SB, Jr., Kristen KH, Trnka HJ. Plantar pressure distribution after tibiotalar arthrodesis. *Clin Biomech (Bristol, Avon)*. Jul 2011;26(6):620-5.
4. Wang Y, Li Z, Wong DW, Zhang M. Effects of Ankle Arthrodesis on Biomechanical Performance of the Entire Foot. *PLoS One*. 2015;10(7):e0134340.
5. Sheridan BD, Robinson DE, Hubble MJ, Winson IG. Ankle arthrodesis and its relationship to ipsilateral arthritis of the hind- and mid-foot. *J Bone Joint Surg Br*. Feb 2006;88(2):206-7.
6. Zwipp H, Rammelt S, Endres T, Heineck J. High union rates and function scores at midterm followup with ankle arthrodesis using a four screw technique. *Clin Orthop Relat Res*. Apr 2010;468(4):958-68.
7. Coester LM, Saltzman CL, Leupold J, Pontarelli W. Long-term results following ankle arthrodesis for post-traumatic arthritis. *J Bone Joint Surg Am*. Feb 2001;83(2):219-28.
8. Hendrickx RP, Stufkens SA, de Bruijn EE, Sierevelt IN, van Dijk CN, Kerkhoffs GM. Medium- to long-term outcome of ankle arthrodesis. *Foot Ankle Int*. Oct 2011;32(10):940-7.
9. Gaedke IE, Wiebking U, O'Loughlin PF, Krettek C, Gaulke R. Clinical and radiological mid- to long-term outcomes following ankle fusion. Article. *In Vivo*. 2018;32(6):1463-1471. doi:10.21873/invivo.11400
10. Jones CR, Wong E, Applegate GR, Ferkel RD. Arthroscopic Ankle Arthrodesis: A 2-15 Year Follow-up Study. *Arthroscopy*. May 2018;34(5):1641-1649.
11. Fuchs S, Sandmann C, Skwara A, Chylarecki C. Quality of life 20 years after arthrodesis of the ankle. A study of adjacent joints. *J Bone Joint Surg Br*. Sep 2003;85(7):994-8.
12. Vier D, Louis T, Fuchs D, Royer CT, Zide JR, Jaffe DE. Radiographic assessment of the subtalar joint: An evaluation of the Kellgren-Lawrence scale and proposal of a novel scale. *Clin Imaging*. Mar 2020;60(1):62-66.
13. Mayich DJ, Pinsker E, Mayich MS, Mak W, Daniels TR. An analysis of the use of the Kellgren and Lawrence grading system to evaluate peritalar arthritis following total ankle arthroplasty. *Foot Ankle Int*. Nov 2013;34(11):1508-15.
14. Cohen MM, Vela ND, Levine JE, Barnoy EA. Validating a new computed tomography atlas for grading ankle osteoarthritis. *J Foot Ankle Surg*. Mar-Apr 2015;54(2):207-13.
15. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. *Ann Rheum Dis*. Dec 1957;16(4):494-502.
16. Aaronson NK, Muller M, Cohen PDA, et al. Translation, Validation, and Norming of the Dutch Language Version of the SF-36 Health Survey in Community and Chronic Disease Populations. *Journal of Clinical Epidemiology*. 1998/11/01/ 1998;51(11):1055-1068. doi:https://doi.org/10.1016/S0895-4356(98)00097-3

17. Sierevelt IN, Beimers L, van Bergen CJA, Haverkamp D, Terwee CB, Kerkhoffs G. Validation of the Dutch language version of the Foot and Ankle Outcome Score. *Knee Surg Sports Traumatol Arthrosc*. Aug 2015;23(8):2413-2419.

18. de Boer AS, Tjioe RJC, Van der Sijde F, et al. The American Orthopaedic Foot and Ankle Society Ankle-Hindfoot Scale; translation and validation of the Dutch language version for ankle fractures. *BMJ Open*. Aug 3 2017;7(8):e017040.

19. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. Mar 1977;33(1):159-74.

20. Habibzadeh F, Habibzadeh P, Yadollahie M. On determining the most appropriate test cut-off value: the case of tests with continuous results. *Biochem Med (Zagreb)*. Oct 15 2016;26(3):297-307.

21. Fluss R, Faraggi D, Reiser B. Estimation of the Youden Index and its associated cutoff point. *Biom J*. Aug 2005;47(4):458-72.

22. Morrey BF, Wiedeman GP, Jr. Complications and long-term results of ankle arthrodeses following trauma. *J Bone Joint Surg Am*. Jul 1980;62(5):777-84.

23. Thomas R, Daniels TR, Parker K. Gait analysis and functional outcomes following ankle arthrodesis for isolated ankle arthritis. *J Bone Joint Surg Am*. Mar 2006;88(3):526-35.

24. Morasiewicz P, Dejnek M, Orzechowski W, Szelerski Ł. Subjective and Objective Outcomes of Ankle Joint Arthrodesis with Either Ilizarov or Internal Fixation. *J Foot Ankle Surg*. Apr 1 2022;

25. Conti MS, Ellis SJ. Weight-bearing CT Scans in Foot and Ankle Surgery. *J Am Acad Orthop Surg*. Jul 15 2020;28(14):e595-e603.

Chapter 4

Bone union assessment with computed tomography (CT) and statistical associations with mechanical or histological testing: A systematic review of animal studies.

A. Willems, C. Içli, J.H. Waarsing, S.M.A. Bierma-Zeinstra, D.E. Meuffels
Calcif Tissue Int. 2022 Feb;110(2):147-161

Abstract

Aim: Objective and accurate assessment of bone union after a fracture, arthrodesis or osteotomy is relevant for scientific and clinical purposes. Bone union is most accurately imaged with computed tomography (CT), but no consensus exists about objective assessment of bone union from CT images. It is unclear which CT-generated parameters are most suitable for bone union assessment. The aim of this review of animal studies is to find which CT-generated parameters are associated most strongly with actual bone union.

Animals and methods: Scientific databases were systematically searched. Eligible studies were studies that 1) were animal studies, 2) created a fracture, 3) assessed bone union with CT, 4) performed mechanical or histological testing as measure of actual bone union, 5) associated CT-generated outcomes to mechanical or histological testing results. Two authors selected eligible studies and performed risk of bias assessment with QUADAS-2 tool.

Results: From 2567 studies that were screened, thirteen studies were included. Most common CT-parameters that were investigated were bone mineral density, bone volume and total callus volume. Studies showed conflicting results concerning the associations of these parameters with actual bone union. CT-assessed torsional rigidity (assessed by three studies) and callus density (assessed by two studies) showed best results. The studies investigating these two parameters reported moderate to strong associations with actual bone union.

Conclusion: CT-assessed torsional rigidity and callus density seem the most promising parameters to represent actual bone union after a fracture, arthrodesis or osteotomy.

Introduction

Achieving bone union is the main goal in patients after a fracture, osteotomy or arthrodesis. But when has bone healed? This is a simple question, but the answer is rather complicated.

In the clinic, bone union is generally assessed based on conventional radiographs and on clinical examination, such as: response to weight bearing or palpation of the fracture [1]. However, assessing bone union is a rather subjective decision [2], and the lack of consensus has been extensively described by several studies [3, 4].

Assessment of bone union after a fracture, arthrodesis or osteotomy is an important clinical consideration. Wrong assessment of bone healing can have major negative consequences for a patient. By overestimating the amount of bone healing, a bone might be loaded too early resulting in a displaced fracture or failure of osteosynthesis material. Underestimating bone healing may cause unnecessary immobilization resulting in stiffness, decreased muscle mass and function, and productivity loss of the patient [5, 6]. Especially if bone union is doubtful, an objective an accurate assessment tool can be helpful in clinical decision making. Also, for scientific purposes an objective and accurate method of fusion assessment would be of high value. Being able to accurately assess bone union would have several advantages like a decreased risk of biases within studies and less patients needed in clinical trials with bone union as primary outcome. Additionally, it would become easier to compare results between studies. In orthopedic studies, bone union is a commonly used primary outcome, for instance in studies investigating bone healing stimulating therapies after a fracture, osteotomy or arthrodesis [7-9]. For the objective assessment of bone healing from radiographs, the radiographic union score (RUS) has been introduced in 2012 [10, 11]. Ever since, this semi quantitative assessment tool for assessment of fracture healing has become increasingly popular as an outcome measure in clinical studies[12, 13]. However, computed tomography (CT) is the best method to image bone, and has been shown to be superior to plain radiographs, MRI and DEXA to assess bone union [14-17]. For CT no golden-standard exists for the objective assessment of bone union as an outcome measure. Therefore, we would like to create a method to objectively assess bone union from CT. This could then be used as golden standard for bone union assessment in clinical studies, but could also be used in the clinic if bone fusion after fracture, arthrodesis or osteotomy is doubtful.

To establish an objective clinically applicable tool for bone union assessment, we need to know which CT-generated outcomes have a strong association with actual bone union. This review will therefore investigate which CT parameters are associated with actual

bone healing. Actual bone union will be tested by mechanical or histological tests. As it is unethical and therefore impossible to acquire this data in clinical studies, in this review we focus on animal studies. The aim of this review is to find CT-parameters that best represent actual bone union, which is indicated by mechanical or histological testing.

Method

The protocol of this review has been prospectively registered at the International prospective register of systematic reviews (http://www.crd.york.ac.uk/prospero/; registration number CRD42020164733).

To find all studies concerning the assessment of bone union with CT, an online search was performed on the 5th of February 2020. Five online databases were searched (Embase.com, Medline Ovid, Web of science, Cochrane CENTRAL, and Google Scholar). The search strategy for Medline Ovid is presented in table 1, and was adapted for the other databases. Following the selection of eligible articles, reference lists of eligible articles were checked for missed articles.

After the search of the databases, eligible articles were selected, by two authors (AW and CI), based on predefined eligibility criteria (table 2). Overall, we included studies that created a fracture in the appendicular skeleton of an animal. A fracture was defined as a bone gap that was created by performing an osteotomy or by impact loading. Studies with distraction osteogenisis or bony defects were excluded. Bony defects were defined as drilling a hole in a bone. After at least 4 weeks, CT should be performed to assess bone union. The time period of 4 weeks was chosen because we aim to look at more advanced fracture healing, and are not interested in the very early stages of bone healing. Simultaneously with CT, actual bone union should be tested by mechanical or histological testing. Parameters that are obtained from mechanical or histological testing and reflect bone union could be for instance, bone mineral density, bone volume or cross-sectional area. The association between CT outcomes and mechanical or histological outcomes should thereafter be statistically examined.

Table 1. Search strategy for Medline Ovid

(fracture healing/ OR Fractures, Ununited/ OR (((bone* OR fracture* OR arthrodes* OR osteotom* OR scaphoid* OR osseous OR bony) ADJ6 (healing OR union* OR nonunion* OR united OR ununited OR consolidation))).ab,ti.) AND (exp "Validation Studies"/ OR "Comparative Study"/ OR exp "psychometrics"/ OR "outcome assessment (health care)"/ OR exp "observer variation"/ OR exp "Health Status Indicators"/ OR exp "reproducibility of results"/ OR exp "discriminant analysis"/ OR (psychometr* OR clinimetr* OR clinometr* OR (outcome ADJ3 (assessment* OR measure*)) OR (observer* ADJ3 variation*) OR ((reproducib* OR reliab* OR unreliab* OR valid* OR coefficient OR homogeneity OR homogeneous OR generaliza* OR generalisa* OR concordance OR repeatab* OR discriminative OR known group OR subscale* OR sensitiv* OR responsive* OR error OR errors) ADJ6 (diagnos* OR observ* OR tomograph* OR radiodiagnos* OR radiograph* OR x-ray*)) OR ((dimension*) ADJ6 (diagnos* OR observ* OR tomograph* OR radiodiagnos* OR radiograph* OR x-ray*) NOT (3-dimension* OR three-dimension*)) OR (internal* ADJ3 consisten*) OR (cronbach* ADJ3 (alpha OR alphas)) OR (item ADJ3 (correlation* OR selection* OR reduction*)) OR agreement OR precision OR imprecision OR (precise* ADJ3 value*) OR (test ADJ3 retest) OR (reliab* ADJ3 (test OR retest)) OR interrater* OR inter-rater* OR intrarater* OR intra-rater* OR intertester* OR inter-tester* OR intratester* OR intra-tester* OR interobserver* OR inter-observer* OR intraobserver* OR intra-observer* OR intertechnician* OR inter-technician* OR intratechnician* OR intra-technician* OR interexaminer* OR inter-examiner* OR intraexaminer* OR intra-examiner* OR interassay* OR inter-assay* OR intraassay* OR intra-assay* OR interindividual* OR inter-individual* OR intraindividual* OR intra-individual* OR interparticipant* OR inter-participant* OR intraparticipant* OR intra-participant* OR kappa OR kappa-s OR kappas OR ((replicab* OR repeated) ADJ3 (measure OR measures OR findings OR result OR results OR test OR tests)) OR (intraclass ADJ3 correlation*) OR (factor ADJ (analys* OR structure*)) OR (multitrait ADJ3 scaling ADJ3 (analysis OR analyses)) OR item discriminant OR (interscale ADJ3 correlation*) OR ((individual OR interval OR rate OR analysis OR values) ADJ3 variabil*) OR (uncertainty ADJ3 (measurement OR measuring)) OR standard error of measurement OR (limit ADJ3 detection) OR minimal detectable concentration OR interpretab* OR ((minimal OR minimally OR clinical OR clinically) ADJ3 (important OR significant OR detectable) ADJ3 (change OR difference)) OR (small* ADJ3 (real OR detectable) ADJ3 (change OR difference)) OR meaningful change OR ceiling effect OR floor effect OR Item response model OR Rasch OR Differential item functioning OR computer adaptive testing OR item bank OR cross-cultural equivalence OR ((defin* OR assess*) ADJ3 quanti*) OR (classif* ADJ3 (union OR consolidat*)) OR (union ADJ3 Score*)).ab,ti.) AND (exp Tomography, X-Ray Computed/ OR exp radiography/ OR Arthrography/ OR Diagnostic Imaging/ OR X ray film/ OR exp radiologists/ OR ((compute* ADJ3 tomograph*) OR radiodiagnos* OR radiolog* OR radiograph* OR x-ray* OR ct OR (cat ADJ (scan*)) OR rontgen* OR roentgen* OR microCT OR ((bone* OR diagnos*) ADJ3 imaging)).ab,ti.) NOT (letter OR news OR comment OR editorial OR congresses OR abstracts).pt. NOT (case reports/ OR case report.ti.)

Firstly, based on the predefined in- and exclusion criteria, the eligibility of studies was assessed by reading title and abstract. Secondly, both authors read the full text of the pre-selected studies and assessed eligibility. After the first and second round, the study selection of both authors was compared. In case of disagreements, a third reviewer decided (DM).

Table 2. In- and exclusion criteria

Inclusion	Exclusion
• Animal study • Bone fracture of the appendicular skeleton • Aim of the study to quantify bony union with micro CT, quantitative CT, multidetector CT, cone beam CT or clinical CT The relation between CT and histological or mechanical testing is statistically assessed Article in English, Spanish, German or Dutch	• Bony defects or Distraction osteogenesis • Follow-up period <4weeks • Data have been published before Review article Full text not available

Data were extracted from eligible studies using a predefined data extraction sheet. Data extraction was performed by one reviewer (AW) and checked by a second reviewer (CI). Disagreements were resolved by reaching consensus. Data that were extracted from the studies were data related to the methodology of the studies (fracture site, number of animals, animal species, use of bone growth stimulating injections, time till CT, type of CT, CT settings, volume of interest, threshold for bone, performance of histological testing and mechanical testing, mechanical test that was performed), outcome measures (outcomes of mechanical or histological testing, and outcomes of CT), and statistical associations between CT-outcomes and mechanical or histological testing.

Risk of bias assessment was done with the QUADAS-2 tool [18], which is a tool for diagnostic studies. Although the tool was originally designed for human studies, we chose this tool because it is the best available tool to assess risk of bias for studies in this review. The risk of bias assessment was done by two authors (AW and CI), and discrepancies were resolved by reaching consensus.

The primary outcome of this systematic review will be the strength of the associations between CT-assessed outcomes and mechanical or histological tested bone union. These associations can be expressed as Pearson's correlation coefficients, coefficients of determination or strength of association in a regression model. To improve readability of this review, all linear Pearson's correlation coefficients will be squared, resulting in coefficients of determination. To distinguish between weak and strong relations, coefficients of determination will be classified as weak (R^2< 0.4), moderate (R^2= 0.4 – 0.7), and strong (R^2>0.7).

Results

The search initially resulted in 5159 studies. After removing the duplicates 2567 studies were screened on title and abstract, resulting in 38 potentially eligible studies. After reading the full-text of those studies, thirteen studies were included in our systematic review (figure 1).

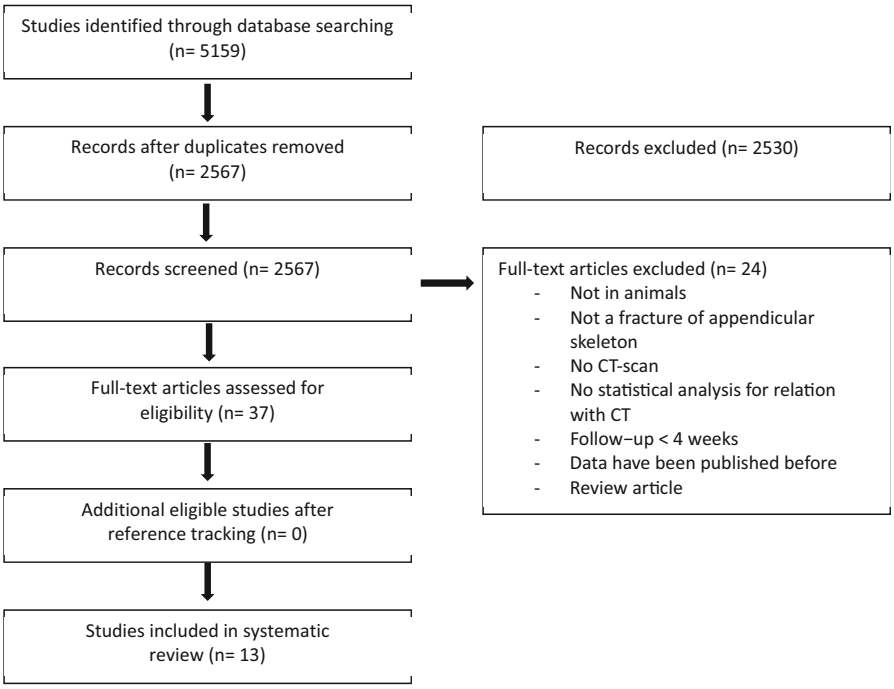


Figure 1 Flow chart of study selection

The results of the risk of bias assessment with the QUADAS-2 tool are presented in table 3. The assessment showed, that risk of bias is generally low in the domains 'animal selection' and 'flow and timing'. However, twelve studies did not clearly describe whether results of the index test (CT) were interpreted without the knowledge of the results of the reference test (mechanical or histological testing) and vice versa. Therefore, the risk of bias concerning these domains is unclear.

Table 3. Risk of bias assessment with the QUADAS-2 tool

Study	RISK OF BIAS				APPLICABILITY CONCERNS		
	Animal selection	Index test	Reference standard	Flow and timing	Animal selection	Index test	Reference standard
Mehta 2013	☺	☺	☺	☺	☺	☺	☺
Morgan 2009	☺	?	?	☺	☺	☺	☺
Nyman 2009	☺	?	?	☺	☺	☺	☺
Shefelbine 2005	☺	?	?	☺	☺	☺	☺
Nazarian 2010	☺	?	?	☺	☺	☺	☺
Fiset 2018	☺	?	?	☺	☺	☺	☺
Jämsä 2000	☺	?	?	☺	☺	☺	☺
Sigurdsen 2011	☺	?	?	☺	☺	☺	☺
Markel 1990	☺	?	?	☺	☺	☺	☺
Augat 1997	☺	?	?	☺	☺	☺	☺
Den Boer 1998	☺	?	?	☺	☺	☺	☺
Wright 2012	☺	?	?	☺	☺	☺	☺
Böhm 1999	☺	☹	?	☺	☺	☺	☺

☺Low risk, ☹High risk, ? Unclear risk

General study characteristics

The studies created a fracture by performing an osteotomy (eight times) [17, 19-25] or by impact loading (five times)[26-30]. Six studies created the fracture in the femur [19, 22, 26, 27, 29, 30], six in the tibia [17, 20, 21, 24, 25, 28], and one in the metatarsus [23]. During follow-up, eight studies used micro CT for the assessment of fracture healing [19, 22, 24-27, 29, 30], two studies peripheral quantitative CT[20, 21], and three studies (quantitative) clinical CT [17, 23, 28]. All studies performed mechanical testing, such as torsional tests [17, 19, 22, 24, 27-30], three point bending tests [20, 23, 26] or axial tests [17, 21]. Two studies also performed histological testing [20], but one of those did not correlate the outcomes to CT outcomes [17]. See table 4 for animal species that were used and more study characteristics.

Linear relations between CT parameters and mechanical or histological outcomes were tested by performing Pearson's correlation [21, 22, 25, 26, 29, 30], bivariate linear regression [17, 19, 20, 23, 24, 28] or multiple regression analysis [27, 29]. Böhm and Jungkunz (1999) also performed bivariate quadratic regression analysis [23].

Table 4. General study characteristics

Study	Fractured bone	Number of animals	Animal species	Bone stimulating injection	Time till CT (weeks)	Type of CT	Histological testing	Mechanical testing
Mehta 2013	Femur	99	Mice	No	2 - 5	Micro	No	Torsional Testing
Morgan 2009	Femur	72	Mice	Yes	2-7	Micro	No	Torsional testing
Nyman 2009	Femur	53	Rats	Yes	4	Micro	No	Destructive three point bending test
Shefelbine 2005	Femur	50	Rats	No	3, 4	Micro	No	Torsional Testing
Nazarian 2010	Femur	10	Rats	Yes	8	Micro	No	Torsional testing
Fiset 2018	Femur	29	Rats	No	5 - 9 or 17	Micro	No	Torsional testing
Jämsä 2000	Tibia	141	Rats	No	4 or 8	Peripheral quantitative	No	Axial tension (4 weeks); axial compression (after 8 weeks)
Sigurdsen 2011	Tibia	40	Rats	No	4, 8.5	Micro	No	Bending test
Markel 1990	Tibia	32	Dogs	No	2,4,8 or 12	Quantitative	No	Torsional testing and indentation testing
Augat 1997	Tibia	28	Sheep	No	9	Peripheral quantitative	Yes	Nondestructive three point bending test
Den Boer 1998	Tibia	24	Goats	Yes	2, 4 or 6	Axial spiral	No	Torsional testing
Wright 2012	Tibia	10	Mice	No	4	Micro	No	Torsion testing
Böhm 1999	Metatarsus	12	Sheep	No	8	Quantitative	No	Nondestructive three-point bending test

Parameters generated with CT representing bone union

Quantitative CT parameters

Quantitative CT parameters that represent bone union are for example bone mineral density (BMD) and total volume of the callus (TV). Studies created volumes of interests (VOI) around the fracture, in which quantitative CT parameters were assessed. Table 5 shows the volumes of interests, bone thresholds and outcome measures that were reported from CT. Also, it shows the parameters assessed from mechanical and histological testing.

Biomechanical CT parameters

Three studies calculated the polar moment of inertia from CT [23, 27, 30]. Polar moment of inertia represents the resistance of bone to torsion, and is dependent on the shape of the callus relative to the torsion axis. Polar moment of inertia is expressed as m^4.

Three studies calculated torsional rigidity (GJ) of the fracture, based on CT-derived data [19, 24, 30]. GJ is a measure describing resistance of a bone when it is subjected to torsional forces, and is expressed as Nm^2. GJ is calculated from the cross sectional area and CT-assessed bone mineral density. GJ was presented as an average of the entire VOI (GJ_{AVG}) [19, 24, 30], and as the weakest slice of the VOI (GJ_{MIN}) [19, 24]. Shefelbine et al. (2005) [30] also calculated the average bending rigidity.

Table 5. Outcome measures on axial CT cross-sections

Study	Volumes of interest	Voxel size (µm)	Energy settings (kV, mA)	Threshold for bone (mgHA/cm^3)	Mechanical or histological testing outcomes	CT-outcomes
Mehta 2013	Fracture callus (region between the outer boundary of the callus and the periosteal surface of the cortex)	10.5	70, 114 mA	190	Torsional stiffness, peak torque	BMD, TMD, TV, BV, BV/TV, Tb.Th, Tb.N, Tb.Sp, $\sigma_{Tb.Th}$, $\sigma_{Tb.Sp}$, Da, Conn.D, SMI,
Morgan 2009	Area between the outer boundary of the callus and the periosteal surface of the pre-existing cortical bone, located between the proximal and distal boundaries of the callus	12	70, 114 mA	641.9	Torsional strength, angular deformation	TMD, BMC, TV, BV, BV/TV, σ_{TMD}, J
Nyman 2009	Trajectory following the outer contours of the tissue; 3.2 mm above and below fracture line Thin region of the outer cortices of first region of interest (representing bridging of the cortices)	32 32	55 and 145 mA	485	Maximum force, bending stiffness, energy to failure	BMD, $BMD_{Bridging\ Cortices,}$ TV, BV, BV/TV, $BV_{Bridging\ Cortices}$/ $TV_{Bridging\ Cortices}$, Min BA/TA; Min $BA_{Bridging\ Cortices}$/$TA_{Bridging\ Cortices}$, Scaled BV/TV, I
Shefelbine 2005	6.2 mm region around the fracture callus	34	NR	Semiautomatically thresholding to highly mineralized bone, newly mineralized bone and soft tissue	Torsional rigidity	BMD_{min}, BV, CSA (min and max), J, GJ, BJ (min, max, and mean)
Nazarian 2010	Osteotomy site and the adjacent bone	34	55, 145 µA	Assessed with global thresholding procedure	Torsional rigidity, peak torque, polar moment of inertia	GJ (min and mean)
Fiset 2018	1.8 mm length of diaphysis consisting of 250 slices that extend proximally from the most distal slice with disruption of the cortical ring of the femur	8	50, 800µA	786	Torsional stiffness, peak torque, failure angle	BMD, TV, BV, BV/TV
Jämsä 2000	Three consecutive cross-sections with a slice distance of 1 mm. The central slice was at mid-callus	148	NR	0.5 cm^{-1} (total bone) 0.93 cm^{-1} (compact bone)	$Failure\ load_{tension}$, $Failure\ load_{Compression}$	BMD, $BMD_{CompactBone}$, BMC, $BMC_{CompactBone}$, CSA, $CSA_{CompactBone}$

Table 5. continued part 2

Study	Volumes of interest	Voxel size (µm)	Energy settings (kV, mA)	Threshold for bone (mgHA/cm³)	Mechanical or histological testing outcomes	CT-outcomes
Sigurdsen 2011	Narrow region near fracture site and wide region encompassing more of the fracture callus region	50.7	NR	Soft callus: 171-540 Hard callus: 540-1200 Cortical bone: >1200	Bending strength	BMD, TV, BV
Markel 1990	Four regions of interest (periosteal callus, endosteal callus, cortex, osteotomy gap)	NR	NR	NR	Peak torque, torsional stiffness, Angular deformation at failure, Indentation stiffness	$BMD_{osteotomy\ gap}$, $BMD_{periosteal\ callus}$, BMD_{cortex}, $BMD_{endosteal\ callus}$
Augat 1997	Seven consecutive transverse sections of which the central slice was at the plane of the former osteotomy	295	47, 0.3 mA	200 mg/cm³	Flexural rigidity	BMD; BMC; CSA
Den Boer 1998	Callus	NR	120, 165mA	NR	Torsional strength, torsional stiffness	CD, CM
Wright 2012	Callus and fracture surface	8	50, 160µA	200 and 590	Peak torque, torsional stiffness	BMD, TMD, TV, BV, BV/TV, Tb.Th, Tb.N, Tb.Sp, SA, BA/TA, GJ (min., mean and surface)
Böhm 1998	Fracture gap automatically detected by algorithm	295	NR	NR	Bending stiffness, deformation	BMD, CD, BMC, J

BMD: bone mineral density; TMD: tissue mineral density; TV: total callus volume; BV: mineralized callus volume; BV/TV: mineralized fraction of the callus; BMC: Bone mineral content; CD: Callus density; CM; Callus mass; CSA; Cross-sectional area; GJ: Torsional rigidity; J: Polar moment or inertia; I: Moment of inertia; BJ: Bending rigidity; Da: degree of anisotropy; Conn.D: connectivity density; SMI: structure modelling index; Tb,N: strut number; Tb.Th: strut thickness; Tb.Sp: strut separation; SA: Failure surface area BA/TA: minimal bone area per total area; SSI: SSI: strength-strain index; NR: Not reported.

Associations between CT and mechanical or histological testing

Quantitative CT outcomes with mechanical testing

The included studies used several quantitative parameters assessed from CT to represent bone union. The results of the studies are shown in table 6a and 6b.

Ten studies correlated bone mineral density (BMD) to mechanical outcome. Six studies did not find associations with R^2>.40 between BMD and mechanical outcomes [21, 22, 24, 26, 29, 30]. Four studies found moderate to strong associations with BMD [17, 20, 23, 25]. Böhm and Jungkunz (1999) also found strong associations for a quadratic association between BMD and mechanical testing [23].

Callus density (CD) was assessed by two studies, which both reported strong associations between CD and mechanical testing [23, 28].

Tissue mineral density (TMD) was assessed by two studies [24, 29]. One study reported weak associations[29], whereas the other study found moderate associations between TMD and mechanical testing[24].

For bone mineral content (BMC), two studies did not find associations with R^2>.40 [20, 21]. One study reported a strong linear and quadratic association for BMC with mechanical testing [23].

Total callus volume (TV) was assessed by five studies. Three studies reported no or weak associations between TV and mechanical outcomes [25, 26, 29]. Two studies reported moderate associations with mechanical outcomes [22, 24].

Mineralized callus volume (BV) was assessed by six studies. Three studies reported no or weak associations for BV with mechanical outcomes [26, 29, 30]. Three studies reported moderate to strong associations between BV and mechanical outcomes [22, 24, 25].

The mineralized fraction of the callus (BV/TV) was assessed by four studies [22, 24, 26, 29], of which one study found a moderate association [26].

Cross sectional area (CSA) was assessed by three studies, and was not associated with mechanical outcomes [20, 21, 30].

Some studies investigated less common CT outcome parameters [24, 26-29]. From these parameters, associations with mechanical outcomes with R^2>.50 were found for trabecular thickness[24], and amount of bone across the failure surface area [24].

Morgan et al. (2009) and Mehta, Heyland, Toben and Duda (2013) created regression models to associate mechanical outcomes to quantitative CT parameters [27, 29]. For maximum torque, a model with TMD, BMC and σTMD explained 62% of the variation (R^2=0.62), and a model with TMD, BV and σTMD explained 61% (R^2=0.61)[27]. For torsional rigidity, a model with TMD, BMC, BV/TV and σTMD explained 70% of the variation (R^2=0.70) [27]. Torsional stiffness could be predicted with a model containing strut thickness, the standard deviation of the strut separation, and strut number (R^2= .55). Torsional strength could be predicted with BMD or BV/TV, strut thickness, standard deviation or strut separation (R^2=0.57).

Quantitative CT outcomes with histological testing

Augat et al. (1997) was the only study who correlated CT-outcomes to histological outcomes. They reported a moderate association (R^2=.62) between minimal BMD and histologically assessed percentage bone in periosteal callus. A strong association (R^2=.71) was reported between the minimal BMD and histologically assessed percentage bone in fracture gap.

Biomechanical CT outcomes

Polar moment of inertia was assessed by three studies. Two studies found no or weak associations between moment of inertia and mechanical outcome [26, 30]. Böhm and Jungkunz (1999) reported moderate linear and quadratic associations between polar moment of inertia and mechanical testing [23].

Three studies associated CT-assessed torsional rigidity to torsional rigidity assessed by mechanical testing [19, 24, 30]. All three studies reported moderate to strong associations between the average torsional rigidity and mechanical testing results [19, 24, 30].

Shefelbine et al. (2005) reported moderate associations between CT assessed maximum and mean bending rigidity and mechanical outcomes [30].

Table 6a. Coefficients of determination for linear associations between CT-outcome measures and mechanical or histological testing outcomes

	Associations with mechanical or histological testing		
Bone mineral denisty (BMD)			
Mehta 2013	.04[B]	.04[D]	
Nyman 2009			
- *Bridging cortices*	.05[G]	.05[H]	<.01[I]
- *Overall*	.20[G]	.11[H]	.03[I]
Shefelbine 2005			
- *Minimum*	.08[J]		
Fiset 2018	.04[B]	.22[D]	.19[E]
Jämsä 2000			
- *Compact bone*	.09[A1]	.18[A2]	
- *Overall*	.04[A1]	.32[A2]	
Sigurdsen 2011			
- *30 days post fracture*			
○ *External fixator*	<.01[F]		
○ *Intramedullary nail*	.68[F]		
- *60 days post fracture*			
○ *External fixator*	.25[F]		
○ *Intramedullary nail*	.01[F]		
Markel 1990			
- *Fracture gap*	.74[B]	.60[D]	.56[K]
- *Periosteal callus*	.73[B]	.49[D]	.64[K]
- *Cortex*	n.s.[B]	n.s.[D]	n.s.[K]
- *Endeosteal callus*	n.s.[B]	.35[D]	.40[K]
Augat 1997			
- *Minimum*	.70[L]	.62[M]	.71[N]
Wright 2012			
- *0.2 gHa/CM3*	n.s.[B]	n.s.[D]	
- *0.59 gHa/CM3*	n.s.[B]	n.s.[D]	
Böhm 1999			
- *Hard bone*	.70[H]		
- *Overall*	.75[H]		
Callus denisty (CD)			
Den Boer 1998	.82[B]	.72[C]	
Böhm 1999	.84[H]		
Tissue mineral denisty (TMD)			
Mehta 2013	.31[B]	.28[D]	
Wright 2012			
- *0.20 gHa/CM³*	.60[B]	n.s.[D]	
- *0.59 gHa/CM³*	.63[B]	n.s.[D]	
Bone mineral content (BMC)			
Jämsä 2000			
- *Compact bone*	.01[A1]		
	.32[A2]		
- *Overall*	>.01[A1]		
	.20[A2]		
Augat 1997	.29[L]		
Böhm 1999	.90[H]		

Table 6a. Continued part 2

	Associations with mechanical or histological testing		
Total callus volume (TV)			
Mehta 2013	.07[B]	.08[D]	
Nyman 2009	.07[G]	.18[H]	.05[I]
Fiset 2018	.56[B]	.45[D]	.30[E]
Sigurdsen 2011			
- *0.17-0.54 gHa/CM³; 30 days post fracture*			
○ *External fixator*	.21[F]		
○ *Intramedullary nail*	.26[F]		
- *0.17-0.54 gHa/CM³; 60 days post fracture*			
○ *External fixator*	.10[F]		
○ *Intramedullary nail*	.18[F]		
- *0.54-1.2 gHa/CM³; 30 days post fracture*			
○ *External fixator*	.23[F]		
○ *Intramedullary nail*	.12[F]		
- *0.54-1.2 gHa/CM³; 60 days post fracture*			
○ *External fixator*	.01[F]		
○ *Intramedullary nail*	.01[F]		
Wright 2012			
- *0.20 gHa/CM³*	.54[B]	n.s.[D]	
- *0.59 gHa/CM³*	.54[B]	n.s.[D]	
Mineralized callus volume (BV)			
Mehta 2013	.08[B]	.14[D]	
Nyman 2009	.10[G]	.07[H]	<.01[I]
Shefelbine 2005	<.01[J]		
Fiset 2018	.60[B]	.67[D]	.47[E]
Sigurdsen 2011			
- *30 days post fracture*			
○ *External fixator*	.17[F]		
○ *Intramedullary nail*	.67[F]		
- *60 days post fracture*			
○ *External fixator*	.15[F]		
○ *Intramedullary nail*	.19[F]		
Wright 2012			
- *0.20 gHa/CM³*	.80[B]	.55[D]	
- *0.59 gHa/CM³*	.79[B]	.60[D]	
Mineralized fraction of the callus (BV/TV)			
Mehta 2013	.01[B]	.01[4]	
Nyman 2009			
- *Bridging cortices*	.20[G]	.17[H]	<.01[I]
- *Overall*	<.01[G]	.03[H]	.12[I]
Fiset 2018	.27[B]	.50[D]	.38[E]
Wright 2012			
- *0.20 gHa/CM³*	n.s.[B]	n.s.[D]	
- *0.59 gHa/CM³*	n.s.[B]	n.s.[D]	

Table 6a. Continued part 3

	Associations with mechanical or histological testing		
Cross-sectional area (CSA)			
Shefelbine 2005			
- *Minimum*	<.01[J]		
- *Maximum*	.01[J]		
Jämsä 2000			
- *Compact bone*	.00[A1]		
	.16[A2]		
- *Overall*	<.01[A1]		
	.04[A2]		
Augat 1997	.15[L]		
Callus mass (CM)			
Den Boer 1998	.05[B]	.02[C]	
Trabecular thickness (Tb.Th)			
Mehta 2013	.34[B]	.32[D]	
Wright 2012			
- *0.20 gHa/CM³*	n.s.[B]	n.s.[D]	
- *0.59 gHa/CM³*	.63[B]	.52[D]	
Trabecular number (Tb.N)			
Mehta 2013	.01[B]	<.01[D]	
Wright 2012			
- *0.20 gHa/CM³*	n.s.[B]	n.s.[D]	
- *0.59 gHa/CM³*	n.s.[B]	n.s.[D]	
Trabecular separation (Tb.Sp)			
Mehta 2013	.02[B]	.05[D]	
Wright 2012			
- *0.20 gHa/CM³*	n.s.[B]	n.s.[D]	
- *0.59 gHa/CM³*	n.s.[B]	n.s.[D]	
Standard deviation Tb.Th			
Mehta 2013	.31[B]	.31[D]	
Standard deviation Tb.Sp			
Mehta 2013	.14[B]	.18[D]	
Failure surface area (SA)			
Wright 2012			
- *Bone 0.2 gHa/CM³*	.73[B]	n.s.[D]	
- *Bone 0.59 gHa/CM³*	.62[B]	.59[D]	
- *Total 0.2 gHa/CM³*	n.s.[B]	n.s.[D]	
- *Total 0.59 gHa/CM³*	n.s.[B]	n.s.[D]	
Bone area per total area (BA/TA)			
Nyman 2009			
- *Bridging cortices*	.29[G]	.21[H]	<.01[I]
- *Overall*	<.01[G]	.06[H]	.05[I]
Wright 2012			
- *0.20 gHa/CM³*	n.s.[B]	n.s.[D]	
- *0.59 gHa/CM³*	n.s.[B]	n.s.[D]	
Degree of anisotropy (DA)			
Mehta 2013	.02[B]	.04[D]	
Connectivity density (Conn.D)			
Mehta 2013	.25[B]	.15[D]	

Table 6a. Continued part 4

	Associations with mechanical or histological testing		
Structure modelling index (SMI)			
Mehta 2013	.10[B]	.12[D]	
Polar moment of inertia			
Nyman 2009			
Overall			
Mean	. 08[G]	.10[H]	.00[I]
Min	<.01[G]	.03[H]	<.01[I]
Bridging cortices			
Mean	<.02[G]	.09[H]	.07[I]
Min	<.01[G]	.01[H]	.09[9]
Shefelbine 2005	.04[J]		
Böhm 1999			
Calculated from center of mass	.68[J]		
Calculated from geometric midpoint	.69[J]		
CT-assessed torsional rigidity			
Shefelbine 2005	.48[J]		
Nazarian 2010			
Smallest	.78[D]	.81[J]	
Average	n.s.[D]	.63[J]	
Wright 2012			
Smallest	n.s. [B]	n.s.[D]	
Average	.50[B]	n.s.[D]	
Surface	.66[B]	n.s.[D]	
CT-assessed bending rigidity			
Shefelbine 2005			
Smallest	.49[J]		
Largest	.52[J]		
Mean	.52[J]		

Dark greys indicate strong (R^2>0.7) associations; light greys indicate moderate associations (R^2 0.4-0.7); A1: failure load tension; A2: failure load compression B: Torsional stiffness; C: Torsional strength; D: Peak torque; E: Failure angle; F: Bending strength; G: Maximum force; H: Bending stiffness; I: Energy to failure; J: Torsional rigidity; K: Identation stiffness; L: Flexural rigidity; M: histologically % bone in periosteal callus; N: histologically % bone in fracture gap; T: Coefficient of determinantion for failure load under tension; C: Coefficient of determinantion for failure load under compression; n.s.: Not significant, no coefficient of determination reported

Table 6b. Coefficients of determination for quadratic associations between CT-outcome measures and mechanical or histological testing outcomes

	Outcome measures of mechanical or histological testing
Bone mineral density (BMD)	
Böhm 1999	
- Hard bone	.74[H]
- Overall	.76[H]
Callus denisty (CD)	
Böhm 1999	.85[H]
Bone mineral content (BMC)	
Böhm 1999	.93[H]
Polar moment of inertia	
Böhm 1999	
- Calculated from center of mass	.68[H]
- Calculated from geometric midpoint	.69[H]
Dark greys indicate strong (R^2>0.7) associations; light greys indicate moderate associations (R^2 0.4-0.7) H: Bending stiffness	

Data synthesis

Overall, for two parameters all studies investigating these parameters found moderate or strong associations. These parameters were CD, which was assessed by two studies, and CT-assessed torsional rigidity, which was assessed by three studies. For BMD, TMD, BMC, TV, BV, trabecular thickness and polar moment of inertia, 30%-60% of the studies investigating these parameters found associations. For BV/TV, CSA, trabecular number, trabecular separation, and bone area per total area, less than 30% of the studies found an association for these parameters.

Some parameters were only assessed by one study. From those, CT-assessed bending rigidity and amount of bone across the failure surface area showed moderate to strong associations.

Discussion

We aimed to identify CT-outcome parameters which are associated most strongly with bone union after a fracture. The associations found by the studies are conflicting, with exception for CT-assessed torsional and bending rigidity, and callus density.

CT-assessed torsional rigidity was found to have moderate to strong associations by all three studies that investigated it. Torsional rigidity is calculated from CT acquired data, and is dependent on the callus density, cross-sectional area and the distribution of bone density within the callus [19, 30]. Based on CT, virtual models of the bone are created on which virtual mechanical testing can be performed. From this virtual testing, torsional rigidity is calculated [19, 30, 31]. Average torsional rigidity showed moderate associations with mechanical tests in all three studies[19, 24, 30]. The results of Naziarian et al. (2010) [19] showed that minimum torsional rigidity had a stronger association with mechanical testing than average torsional rigidity. This means that analyzing only the weakest segment (axial slice) of CT images would give the strongest associations. This seems logical, as failure of a beam under forces is dependent on the weakest point, and not the average strength [19]. However, Wright, Nam and Whyne (2012)[24] did not find an association between minimum torsional rigidity and mechanical testing. According to Wright, the use of the tibia, and not the femur as Nazarian did, might explain this [24]. In contrast to the femur, the diameter of the tibia decreases when going more distally. As torsional rigidity is dependent on the CSA, the minimum torsional rigidity might therefore move to the most distal part of the VOI when analyzing the tibia [24]. This once more indicates that the assessment of fracture healing is complex and dependent on many variables.

This complexity may have led to the conflicting results of the other parameters. For example, quite strong association were reported for BMD by three studies, whereas other studies found no associations with BMD. Because of the conflicting results between studies the generalizability of the associations seems to be quite low. Also, most studies in this review explored linear relations, but Böhm and Jungkunz (1999) showed that associations might be quadratic [23]. However, Böhm and Jungkunz (1999) was the only study investigating quadratic associations and it was a small study (n=12).

So far, CT-assessed torsional rigidity seems a promising parameter for bone union assessment. Clinically, several studies have been investigating CT-assessed torsional rigidity. CT-assessed torsional rigidity was successfully used for the prediction of fractures in patients with bone lesions [32-34]. Also, recently the first clinical study has been published that used CT-assessed torsional rigidity to assess tibial fracture healing [35]. In this study, a low dose CT was made of the tibia 12 weeks after surgical fixation. Software was used to create a virtual model of the fractured tibia which was adapted to a model

of an intact tibia. Virtual torsional testing could then be performed on these models, resulting in torsional rigidity values for the fractured and intact tibia. Lastly, torsional rigidity of the fractured model was divided by the torsional rigidity of the intact model. By doing this, a dimensionless parameter was created which indicates the progression of healing relative to the intact tibia [31]. Given the results of this review, and the promising results of the first clinical study, CT-assessed torsional rigidity could become a useful tool for bone union assessment. However, at this moment the clinical applicability of CT-assessed torsional rigidity is limited. Advanced software and knowledge are needed to conduct CT-based structural rigidity analysis (CTRA)[32]. Although CTRA can be done with data from any CT-scanner, bone densities are very important for the analysis. Therefore, phantoms with known bone densities should be scanned with the patient [32].

This systematic review encountered some limitations. Firstly, CT-assessed torsional rigidity and callus density were only assessed by a limited number of studies (three studies for CT-assessed torsional rigidity, and two studies for callus density). Although those studies show promising results, more studies should be done to further confirm these results. Parameters that were assessed by more than three studies, had higher chances of finding contradictory results. However, the more investigated parameters in this review showed no significant associations in most of the studies. BMD for instance was investigated by ten studies, of which only four reported significant associations. A second limitations of this review is that the statistical associations that are presented come from animal studies. We should be careful by translating these results directly to clinical human fractures, as data retrieved from animal studies might be unreliable in clinical studies[10]. For example, studies have shown that rodent bone remodeling is different from large animal or human bone remodeling, because it is lacking intracortical remodeling[36]. Therefore, associations for bone healing might be different for rodents compared to large animals or humans. Also, most studies in this review used micro CT-scanners with higher spatial resolutions and higher radiation doses than clinical CT-scanners [19, 21, 26, 27]. Therefore, clinical CT-scanners might be less accurate than micro CT-scanners [37]. , Thirdly, for this systematic review we used fairly strict inclusion criteria. The main reason for these strict criteria was to keep heterogeneity between studies as low as possible to be able to compare studies, and therewith draw a firm conclusion. Even with these strict criteria, the heterogeneity between studies was high. Studies used different location of fractures, animal species, scanning protocols and mechanical testing protocols, which is likely to affect the associations found between the studies. Also, four studies used drug treatments to increase fracture healing[19, 26-28]. These treatments can modulate structural and mechanical properties of the callus[38]. Due to the strict inclusion criteria, many studies were excluded during the study selection process. These were also studies who assessed bone healing by performing CT, mechanical and histological

testing. However, in these studies the different methods were used complementary to each other and the results of these methods were not compared to each other. Therefore it is not possible to draw conclusions from these studies concerning the best CT outcome parameter. Also, minimal follow-up time was set to 4 weeks, as we were not interested in studies who only looked at early stages of fracture healing. As fracture healing progresses differently between animal species and depends on fracture size, on could argue if this period was accurate. Also, the associations between CT-parameters and mechanical and histological outcomes might be influenced by the stage of fracture healing, which may vary between the studies. Lastly, the risk of bias of studies was assessed with the QUADAS-2 tool. As this tool is designed for clinical studies, it may not be accurate for pre-clinical studies. However, no pre-clinical risk of bias tool exists for diagnostic studies. Most studies in this review showed concerns about risk of bias. To decrease risk of bias in future studies, we strongly recommend to interpret the index test (CT), without knowing the results of the reference test (mechanical or histological testing), and to describe this process in the paper.

Based on the currently available literature, density related parameters seem to be most promising parameters to assess bone union after a fracture. Especially CT-assessed torsional rigidity is a promising parameter to assess bone union. To improve the clinical assessment of fracture healing, we encourage the conduction of more high quality clinical studies investigating the applicability of CT-assessed torsional rigidity for bone union assessment. In the future, torsional rigidity could potentially become a widely accepted outcome measure for bone union assessment in clinical studies and in clinical practice.

References

1. Bhandari M, Fong K, Sprague S, Williams D, Petrisor B (2012) Variability in the definition and perceived causes of delayed unions and nonunions: a cross-sectional, multinational survey of orthopaedic surgeons. J Bone Joint Surg Am 94:e1091-1096. https://doi.org/10.2106/JBJS.K.01344
2. Bhandari M, Guyatt GH, Swiontkowski MF, Tornetta P, 3rd, Sprague S, Schemitsch EH (2002) A lack of consensus in the assessment of fracture healing among orthopaedic surgeons. J Orthop Trauma 16:562-566. https://doi.org/10.1097/00005131-200209000-00004
3. Cunningham BP, Brazina S, Morshed S, Miclau T, 3rd (2017) Fracture healing: A review of clinical, imaging and laboratory diagnostic options. Injury 48 Suppl 1:S69-S75. https://doi.org/10.1016/j.injury.2017.04.020
4. Corrales LA, Morshed S, Bhandari M, Miclau T, 3rd (2008) Variability in the assessment of fracture-healing in orthopaedic trauma studies. J Bone Joint Surg Am 90:1862-1868. https://doi.org/10.2106/JBJS.G.01580
5. Dudley GA, Duvoisin MR, Adams GR, Meyer RA, Belew AH, Buchanan P (1992) Adaptations to unilateral lower limb suspension in humans. Aviat Space Environ Med 63:678-683
6. Bloomfield SA (1997) Changes in musculoskeletal structure and function with prolonged bed rest. Med Sci Sports Exerc 29:197-206. https://doi.org/10.1097/00005768-199702000-00006
7. Willems A, van der Jagt OP, Meuffels DE (2019) Extracorporeal Shock Wave Treatment for Delayed Union and Nonunion Fractures: A Systematic Review. J Orthop Trauma 33:97-103. https://doi.org/10.1097/BOT.0000000000001361
8. Bovbjerg P, Høgh D, Froberg L, Schmal H, Kassem M (2021) Effect of PTH treatment on bone healing in insufficiency fractures of the pelvis: a systematic review. EFORT Open Rev 6:9-14. https://doi.org/10.1302/2058-5241.6.200029
9. Van Lieshout EMM, Den Hartog D (2020) Effect of platelet-rich plasma on fracture healing. Injury. https://doi.org/10.1016/j.injury.2020.12.005
10. Schwarzenberg P, Darwiche S, Yoon RS, Dailey HL (2020) Imaging Modalities to Assess Fracture Healing. Current Osteoporosis Reports 18:169-179. https://doi.org/10.1007/s11914-020-00584-5
11. Whelan DB, Bhandari M, Stephen D, Kreder H, McKee MD, Zdero R, Schemitsch EH (2010) Development of the radiographic union score for tibial fractures for the assessment of tibial fracture healing after intramedullary fixation. J Trauma 68:629-632
12. Leow JM, Clement ND, Simpson A (2020) Application of the Radiographic Union Scale for Tibial fractures (RUST): Assessment of healing rate and time of tibial fractures managed with intramedullary nailing. Orthop Traumatol Surg Res 106:89-93
13. Litrenta J, Tornetta P, 3rd, Mehta S, Jones C, O'Toole RV, Bhandari M, Kottmeier S, Ostrum R, Egol K, Ricci W, Schemitsch E, Horwitz D (2015) Determination of Radiographic Healing: An Assessment of Consistency Using RUST and Modified RUST in Metadiaphyseal Fractures. J Orthop Trauma 29:516-520
14. Braunstein EM, Goldstein SA, Ku J, Smith P, Matthews LS (1986) Computed tomography and plain radiography in experimental fracture healing. Skeletal Radiology 15:27-31. https://doi.org/10.1007/bf00355070

15. Firoozabadi R, Morshed S, Engelke K, Prevrhal S, Fierlinger A, Miclau T, 3rd, Genant HK (2008) Qualitative and quantitative assessment of bone fragility and fracture healing using conventional radiography and advanced imaging technologies--focus on wrist fracture. J Orthop Trauma 22:S83-90. https://doi.org/10.1097/BOT.0b013e31815ea2a4

16. Grigoryan M, Lynch JA, Fierlinger AL, Guermazi A, Fan B, MacLean DB, MacLean A, Genant HK (2003) Quantitative and qualitative assessment of closed fracture healing using computed tomography and conventional radiography. Acad Radiol 10:1267-1273. https://doi.org/10.1016/s1076-6332(03)00467-7

17. Markel MD, Wikenheiser MA, Morin RL, Lewallen DG, Chao EY (1990) Quantification of bone healing. Comparison of QCT, SPA, MRI, and DEXA in dog osteotomies. Acta Orthop Scand 61:487-498. https://doi.org/10.3109/17453679008993569

18. Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MMG, Sterne JAC, Bossuyt PMM, the Q-G (2011) QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies. Annals of Internal Medicine 155:529-536. https://doi.org/10.7326/0003-4819-155-8-201110180-00009

19. Nazarian A, Pezzella L, Tseng A, Baldassarri S, Zurakowski D, Evans CH, Snyder BD (2010) Application of structural rigidity analysis to assess fidelity of healed fractures in rat femurs with critical defects. Calcif Tissue Int 86:397-403. https://doi.org/10.1007/s00223-010-9353-4

20. Augat P, Merk J, Genant HK, Claes L (1997) Quantitative assessment of experimental fracture repair by peripheral computed tomography. Calcif Tissue Int 60:194-199. https://doi.org/10.1007/s002239900213

21. Jamsa T, Koivukangas A, Kippo K, Hannuniemi R, Jalovaara P, Tuukkanen J (2000) Comparison of radiographic and pQCT analyses of healing rat tibial fractures. Calcif Tissue Int 66:288-291. https://doi.org/10.1007/s002230010058

22. Fiset S, Godbout C, Crookshank MC, Zdero R, Nauth A, Schemitsch EH (2018) Experimental validation of the radiographic union score for tibial fractures (RUST) using micro-computed tomography scanning and biomechanical testing in an in-vivo rat model. J Bone Jt Surg Am Vol 100:1871-1878. https://doi.org/10.2106/jbjs.18.00035

23. Martin Böhm A, Jungkunz B (1999) Bending stiffness of healing fractures can be calculated from quantitative computed tomography. Eur J Radiol 30:28-32. https://doi.org/10.1016/s0720-048x(98)00138-7

24. Wright DA, Nam D, Whyne CM (2012) A comparison of stereology, structural rigidity and a novel 3D failure surface analysis method in the assessment of torsional strength and stiffness in a mouse tibia fracture model. J Biomech 45:2236-2240. https://doi.org/10.1016/j.jbiomech.2012.06.016

25. Sigurdsen U, Reikeras O, Hoiseth A, Utvag SE (2011) Correlations between strength and quantitative computed tomography measurement of callus mineralization in experimental tibial fractures. Clin Biomech 26:95-100. https://doi.org/10.1016/j.clinbiomech.2010.09.004

26. Nyman JS, Munoz S, Jadhav S, Mansour A, Yoshii T, Mundy GR, Gutierrez GE (2009) Quantitative measures of femoral fracture repair in rats derived by micro-computed tomography. J Biomech 42:891-897. https://doi.org/10.1016/j.jbiomech.2009.01.016

27. Morgan EF, Mason ZD, Chien KB, Pfeiffer AJ, Barnes GL, Einhorn TA, Gerstenfeld LC (2009) Micro-computed tomography assessment of fracture healing: relationships among callus structure, composition, and mechanical function. Bone 44:335-344. https://doi.org/10.1016/j.bone.2008.10.039

28. den Boer FC, Bramer JAM, Patka P, Bakker FC, Barentsen RH, Feilzer AJ, de Lange ESM, Haarman HJTM (1998) Quantification of fracture healing with three-dimensional computed tomography. Archives of Orthopaedic and Trauma Surgery 117:345-350. https://doi.org/10.1007/s004020050263

29. Mehta M, Heyland M, Toben D, Duda GN (2013) Microstructure and homogeneity of distribution of mineralised struts determine callus strength. Eur Cell Mater 25:366-379; discussion 378-369. https://doi.org/10.22203/ecm.v025a26

30. Shefelbine SJ, Simon U, Claes L, Gold A, Gabet Y, Bab I, Muller R, Augat P (2005) Prediction of fracture callus mechanical properties using micro-CT images and voxel-based finite element analysis. Bone 36:480-488. https://doi.org/10.1016/j.bone.2004.11.007

31. Schwarzenberg P, Maher MM, Harty JA, Dailey HL (2019) Virtual structural analysis of tibial fracture healing from low-dose clinical CT scans. J Biomech 83:49-56. https://doi.org/10.1016/j.jbiomech.2018.11.020

32. Damron TA, Nazarian A, Entezari V, Brown C, Grant W, Calderon N, Zurakowski D, Terek RM, Anderson ME, Cheng EY, Aboulafia AJ, Gebhardt MC, Snyder BD (2016) CT-based Structural Rigidity Analysis Is More Accurate Than Mirels Scoring for Fracture Prediction in Metastatic Femoral Lesions. Clin Orthop Relat Res 474:643-651. https://doi.org/10.1007/s11999-015-4453-0

33. Leong NL, Anderson ME, Gebhardt MC, Snyder BD (2010) Computed tomography-based structural analysis for predicting fracture risk in children with benign skeletal neoplasms: comparison of specificity with that of plain radiographs. J Bone Joint Surg Am 92:1827-1833. https://doi.org/10.2106/JBJS.I.00871

34. Snyder BD, Hauser-Kara DA, Hipp JA, Zurakowski D, Hecht AC, Gebhardt MC (2006) Predicting fracture through benign skeletal lesions with quantitative computed tomography. J Bone Joint Surg Am 88:55-70. https://doi.org/10.2106/JBJS.D.02600

35. Dailey HL, Schwarzenberg P, Daly CJ, Boran SAM, Maher MM, Harty JA (2019) Virtual Mechanical Testing Based on Low-Dose Computed Tomography Scans for Tibial Fracture: A Pilot Study of Prediction of Time to Union and Comparison with Subjective Outcomes Scoring. J Bone Joint Surg Am 101:1193-1202. https://doi.org/10.2106/JBJS.18.01139

36. Allen MR (2017) Preclinical Models for Skeletal Research: How Commonly Used Species Mimic (or Don't) Aspects of Human Bone. Toxicol Pathol 45:851-854

37. Cook SD, Patron LP, Salkeld SL, Smith KE, Whiting B, Barrack RL (2009) Correlation of computed tomography with histology in the assessment of periprosthetic defect healing. Clin Orthop Relat Res 467:3213-3220. https://doi.org/10.1007/s11999-009-1087-0

38. Gabet Y, Müller R, Regev E, Sela J, Shteyer A, Salisbury K, Chorev M, Bab I (2004) Osteogenic growth peptide modulates fracture callus structural and mechanical properties. Bone 35:65-73

Chapter 5

How to assess consolidation after foot and ankle arthrodesis with computed tomography. A systematic review.

A. Willems, C.M. Houkes, S.M.A. Bierma-Zeinstra, D.E. Meuffels
Eur J Radiol. 2022 Sep 5;156:110511

Abstract

Purposes: Many studies have been performed that investigate consolidation after arthrodesis of foot and ankle joints. Consolidation in foot and ankle joints is best assessed by computed tomography (CT). However, no golden-standard methodology exists for radiological consolidation assessment from CT after ankle and foot arthrodesis. The aim of this review is to present an overview of the radiological methodologies for consolidation assessment, outcomes on reliability and validity and to advise which methodology should be used

Method: Scientific databases were systematically searched. Eligible studies were studies that 1) performed foot or ankle arthrodesis, 2) mentioned radiological or CT follow-up in abstract, 3) performed postoperative CT in > 50% of patients. Two authors selected eligible studies and performed a risk of bias assessment with the COSMIN tool.

Results: Risk of bias assessment showed that most studies (80%) were at high risk of bias due to poor methodology. The most popular method for consolidation assessment is by subjectively categorizing consolidation into consolidation groups, with a substantial reliability score. Another popular method is to calculate the fusion ratio and then apply a fusion threshold, to distinguish between fused and non-fused joints. This method had an excellent reliability score. In most studies a fusion threshold of 50% is used. However, four studies in this review showed that a 30% fusion threshold may by more valid.

Conclusion: Based on the results of this review we would advise to calculate fusion threshold and apply a 30% fusion threshold to distinguish fused from non-fused foot and ankle joints.

Keywords: Consolidation; Fusion; Ankle; Foot; Arthrodesis; Computed tomography

Introduction

Arthrodesis of foot or ankle joints can be indicated as a result of severe osteoarthritis, malformation or fractures. A dreaded complication after arthrodesis is nonunion of the fused joint. In the literature wide ranges of union rates after arthrodesis of foot or ankle joints have been reported, ranging from 64% till 100% [1]. Risk factors for nonunion have been studied extensively, as have operative and fixation techniques, all to prevent nonunion after arthrodesis [1, 2]. However, although consolidation is a very important outcome measure in these studies, there is no general consensus on how to radiologically assess consolidation after arthrodesis. Due to the lack of consensus, studies currently use different methods to radiologically assess consolidation after arthrodesis.

Little is known about valid and reliable radiological assessment of bone consolidation after foot or ankle arthrodesis. Assessment of bone healing after a fracture is a somewhat more studied subject. A systematic review from Corrales et al. (2008), showed a lack of consensus on fracture healing assessment. They reported that in 123 studies that investigated therapeutic clinical effects in long-bone fractures, 11 different criteria for fracture healing were used. The majority of studies assessed fracture healing with a combination of clinical and radiographic criteria (62%). Clinical criteria for fracture union are pain on weight bearing and pain on palpation. Radiographic criteria for fracture union were bridging of the fracture by callus, trabeculae or osseous bone and obliteration of the fracture line [3]. Only two of 123 studies reported a quantitative measure of the reliability of the radiographic assessment of fracture healing. After the review of Corrales et al. (2008), the 'Radiographic Union Score for Tibial fractures' (RUST) was introduced as a tool for tibial fracture healing assessment. In this tool, anteroposterior and lateral radiographs are scored based on presence or absence of fracture line and callus, resulting in scores ranging between 4 (no union) to 12 (complete healing) [4]. Also, the interobserver agreement for RUST tool were high (>.80) [4]. Cooke et al. (2018) correlated the outcome of RUST to biomechanical properties in femoral mice fractures, and reported high correlations between RUST scores and physical properties [5]. Nowadays, RUST and a slightly adapted version (modified RUST) are commonly used tools for fracture healing assessment [5-8].

For studies that study consolidation after foot or ankle arthrodesis, it would be of great benefit if a validated and widely accepted tool would be used. There is no general consensus on a validated golden-standard for radiological consolidation assessment after arthrodesis. Coughlin et al. (2006) showed however that progress of consolidation of the hindfoot, cannot be accurately assessed from radiographs. Therefore, for consolidation assessment after hindfoot arthrodesis, CT is strongly recommended [9, 10]. In order to get

closer to consensus on consolidation assessment in foot and ankle joints, we should first investigate which methods are currently used for consolidation assessment. Therefore, the aim of this systematic review is to present how studies assess bone consolidation after arthrodesis of foot or ankle joints from CT.

Method

The protocol of this study has been registered (https://www.crd.york.ac.uk/prospero/; registration number: CRD42021199088).

We selected studies which preformed arthrodesis of the foot or ankle joints, which evaluated consolidation with CT. Inclusion criteria were: foot or ankle arthrodesis was performed, mentioning in the abstract of radiological or CT follow-up, postoperative CT to assess bone consolidation is performed in at least 50% of study patients, follow-up period of at least four weeks, studies should be written in English, Dutch, German or Spanish. We excluded reviews and case reports (≤5 participants), animal studies and studies without access to full-text.

An extensive electronic literature search was performed on the 10th of September 2020. Databases that were searched were: Embase, Medline Ovid, Web of science, Cochrane central. See table 1 for the search strategy for Embase. To select eligible studies from the search results, two rounds of study selection were performed. Each round was performed by two independent reviewers (AW and CH). In the first round, search results were screened for eligibility based on title and abstract. Results between reviewers were compared and differences were resolved by consensus. In the second round, the studies selected by the first round were further assessed for eligibility by reading the full text. Results between the reviewers were again compared and resolved by consensus. In case of disagreements, a third reviewer (DM) was consulted.

Table 1. Search strategy for Embase

('arthrodesis'/exp OR (arthrodes* OR spondylodesis* OR ((spin* OR joint* OR ankle* OR hindfoot* OR forefoot* OR foot OR feet OR tarsal* OR metatarsal* OR subtalar* OR talocrural* OR wrist OR knee OR lumbar* OR thora* OR cervi* OR vertebra* OR Occipit* OR C1 OR C2 OR C3 OR C4 OR C5 OR C6 OR C7 OR L1 OR L2 OR L3 OR L4 OR L5 OR T1 OR T2 OR T3 OR T4 OR T5 OR T6 OR T7 OR T8 OR T9 OR T10 OR T11 OR T12) NEAR/6 fusion*)):ab,ti) AND ('fracture healing'/de OR 'fracture nonunion'/de OR 'callus'/de OR 'bone development'/de OR 'ossification'/de OR 'osteoclastogenesis'/de OR 'fusion rate'/de OR (((bone* OR fracture* OR arthrodes* OR osteotom* OR scaphoid* OR osseous OR bony OR spin* OR fusion*) NEAR/6 (healing OR union* OR nonunion* OR united OR ununited OR consolidation OR development* OR formation* OR bridging)) OR (joint* NEAR/3 fusion*) OR callus OR ossificatio* OR osteoclastogenes* OR fusion-rate* OR fusion-status*):ab,ti) AND ('computer assisted tomography'/exp OR 'radiodiagnosis'/de OR radiography/de OR 'bone radiography'/exp OR 'pelvis radiography'/de OR 'joint radiography'/exp OR 'radiological parameters'/de OR radiologist/de OR 'diagnostic imaging'/de OR ((compute* NEAR/3 tomograph*) OR radiodiagnos* OR radiolog* OR radiograph* OR ct OR (cat NEXT/1 (scan*)) OR ((bone* OR diagnos*) NEAR/3 imaging)):ab,ti) NOT ([Conference Abstract]/lim OR [Letter]/lim OR [Note]/lim OR [Editorial]/lim) NOT ('case reports'/de OR 'case report':ti) NOT ([animals]/lim NOT [humans]/lim)

Data were extracted with a pre-defined data extraction sheet. General study data that were extracted were: study type, number of patients, patient population, CT settings, primary aim of the study and if consolidation was the primary outcome of the study. Furthermore, methodological descriptions of consolidation assessment were extracted, who assessed consolidation (radiologist, orthopedic surgeon), how many assessors assessed consolidation, if assessors were blinded, reliability measures (inter- or intraclass correlation coefficient (ICC) or Kappa scores), validity measures (correlations with other methods of assessment or clinical outcomes), and if a study referred to methodologies of other studies.

Risk of bias assessment was assessed with the COSMIN tool which is advised by Cochrane for clinimetric studies [11]. This tool is established to assess the quality of studies on reliability measurements. The tool consists of nine subdomains which are scored as very good, adequate, doubtful, inadequate or not applicable. Five of those subdomains are about design requirements, like time interval between measurements, measurement conditions and blinding of assessors. Three subdomains are about the statistical method used to assess reliability and assess whether an ICC or kappa score was calculated. One subdomain asks whether there are 'any other important flaws in the design or statistical methods of the study'. As we did not only include clinimetric studies, some studies referred to another study for the methodology of consolidation assessment. If so, the quality of this study was assessed with the COSMIN tool and the outcome was included in this subdomain. The risk of bias is assessed by applying the worst-score-count method [11]. The assessment was performed by two reviewers, discrepancies between results were resolved by consensus.

A table will be presented with all included studies and the methods of consolidation assessment from CT. If available, reliability scores of those methods will be presented. They will be interpreted according to the Koch-Landis interpretation of Kappa tests (slight agreement (Kappa=.01-.20), fair agreement (Kappa=.21-.40), moderate agreement (Kappa=.41-.60), substantial agreement (Kappa=.61-.80) and excellent agreement (Kappa=.81-1.00)) [12].

Results

Our electronic search resulted in 12.058 studies. After removing duplicates 6450 studies remained for screening. Based on title and abstract screening, 6356 studies were excluded. The remaining 94 studies were assessed based on full-text. Full-text screening led to exclusion of 60 studies. Therefore, 34 studies were eligible for inclusion in this systematic review. See figure 1 for the flow-chart of study selection.

Risk of bias assessment

The results of the risk of bias assessment are shown in table 2. The results show that 27 studies were assessed as inadequate [9, 10, 13-37], meaning that there is a high risk of bias. This was a result of absence of reporting reliability measurements, such as ICC or Kappa. Also, none of these studies referred to a study with an adequate methodological description of consolidation assessment. Risk of bias was assessed as doubtful in four studies [38-41]. Three of those studies did assess intraobserver reliability but not interobserver reliability [38-40]. One study did not report intrabovserver reliability but the study referred to a study that assessed intraobserver reliability [41]. Three studies were assessed as adequate [42-44]. These studies assessed kappa score's as a measure for agreement between different observers. However it was not clearly described if the assessors of bone union were independent and blinded [42, 43], or if the assessors were without knowledge of each other's scores [44]. Overall, there is a high risk of bias in most of the studies.

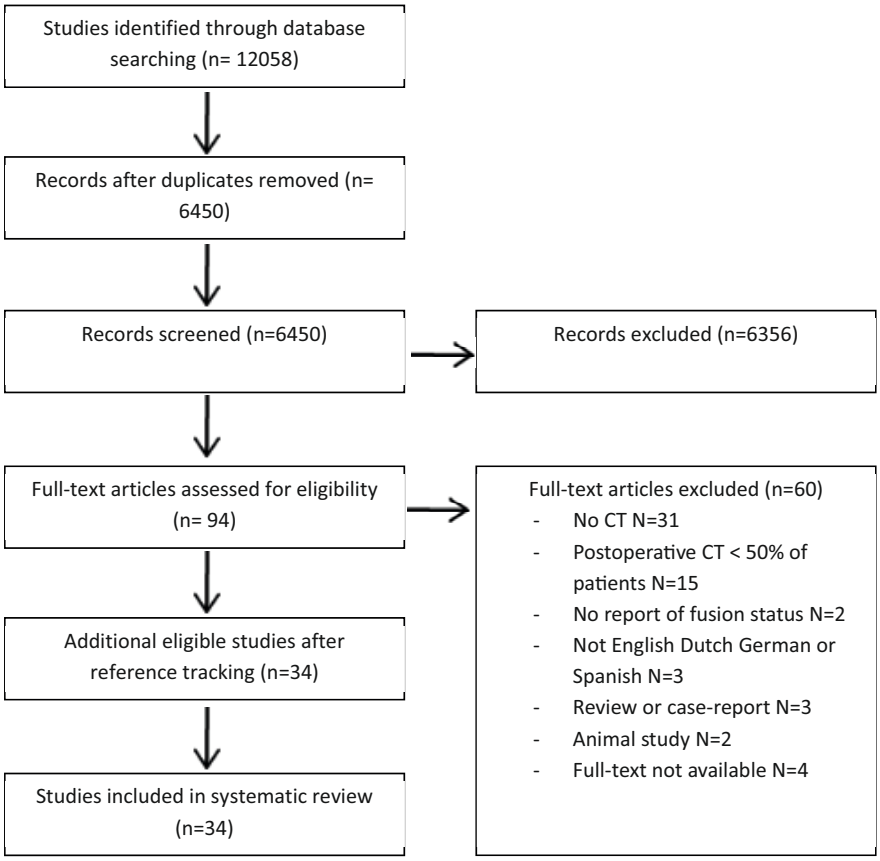


Figure 1. Flow-chart of study selection

General characteristics

In this systematic review, 34 studies were included who assessed consolidation after foot or ankle arthrodesis with CT. The primary outcome in 30 of those studies was assessment of consolidation with CT [9, 10, 13-24, 27-35, 37, 39-42, 44]. Seven of these studies were randomized controlled trials [18, 19, 25, 26, 28, 40, 41], and 27 studies were retro- or prospective cohort studies [9, 10, 13-17, 20-24, 27, 29-39, 42-44].

Table 2. COSMIN risk of bias assessment tool to assess quality of studies on reliability

	Subdomains for scoring Risk of Bias									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**Score**
Jones et al. 2006	N.a.	N.a.	N.a.	N.a.	N.a.	I.ad.	N.a.	N.a.	N.a.	I.ad.
Dorsey et al. 2009	N.a.	N.a.	N.a.	I.ad	N.a.	I.ad.	N.a.	N.a.	N.a.	I.ad.
Thaunat et al. 2012	N.a.	N.a.	N.a.	N.a.	N.a.	I.ad.	N.a.	N.a.	N.a.	I.ad.
Coulomb et al. 2019	N.a.	N.a.	N.a.	N.a.	N.a.	I.ad.	N.a.	N.a.	N.a.	I.ad.
Fourman et al. 2014	N.a.	N.a.	N.a.	N.a.	N.a.	I.ad.	N.a.	N.a.	N.a.	I.ad.
Mehlhorn et al. 2020	N.a.	N.a.	V.g.	Ad.	Ad.	V.g.	N.a.	V.g.	N.a.	Ad.
Coughlin et al. 2006	N.a.	N.a.	N.a.	N.a.	N.a.	I.ad.	N.a.	N.a.	N.a.	I.ad.
Coughlin et al. 2008	N.a.	N.a.	N.a.	N.a.	N.a.	I.ad.	N.a.	N.a.	N.a.	I.ad.
Glazebrook et al. 2013a	N.a.	N.a.	N.a.	N.a.	N.a.	I.ad.	N.a.	N.a.	N.a.	I.ad.
Usuelli et al. 2016	N.a.	N.a.	Ad.	Ad.	Ad.	V.g.	N.a.	N.a.	V.g.	Ad.
Myerson et al. 2019	N.a.	N.a.	N.a.	N.a.	N.a.	I.ad.	N.a.	N.a.	N.a.	I.ad.
Ford et al. 2019	N.a.	N.a.	Df.	Df.	Df.	I.ad.	N.a.	N.a.	N.a.	I.ad.
Bejarano-Pineda 2020	N.a.	N.a.	N.a.	N.a.	N.a.	I.ad.	N.a.	N.a.	N.a.	I.ad.
Wanivenhaus et al. 2017	N.a.	N.a.	N.a.	N.a.	N.a.	I.ad.	N.a.	N.a.	N.a.	I.ad.
Jones et al. 2015	N.a.	N.a.	N.a.	N.a.	N.a.	I.ad.	N.a.	N.a.	N.a.	I.ad.
Daniels 2010	N.a.	N.a.	N.a.	N.a.	N.a.	I.ad.	N.a.	N.a.	N.a.	I.ad.
DiGiovanni et al. 2011	N.a.	N.a.	N.a.	N.a.	N.a.	I.ad.	N.a.	N.a.	N.a.	I.ad.
Glazebrook et al. 2013b	N.a.	N.a.	N.a.	N.a.	N.a.	I.ad.	N.a.	N.a.	N.a.	I.ad.
Dekker et al. 2018	N.a.	N.a.	Df.	Df.	Df.	I.ad.	N.a.	N.a.	N.a.	I.ad.
DiGiovanni et al. 2013	N.a.	V.g	Ad.	Ad.	Ad.	Df.	N.a.	V.g.	N.a.	Df.
Daniels et al. 2019	N.a.	N.a.	N.a.	N.a.	N.a.	Df.	N.a.	N.a.	N.a.	Df.
Daniels et al. 2015	N.a.	N.a.	N.a.	N.a.	N.a.	I.ad.	N.a.	N.a.	N.a.	I.ad.
Kodama et al. 2016	N.a.	N.a.	N.a.	N.a.	N.a.	I.ad.	N.a.	N.a.	N.a.	I.ad.
Aubret et al. 2018	N.a.	N.a.	N.a.	N.a.	N.a.	I.ad.	N.a.	N.a.	N.a.	I.ad.
Bibbo 2009	N.a.	N.a.	N.a.	N.a.	N.a.	I.ad.	N.a.	N.a.	N.a.	I.ad.
Cerrato et al. 2014	N.a.	N.a.	Ad.	Ad.	Ad.	N.a.	N.a.	N.a.	V.g.	Ad.
DiGiovanni et al. 2016	N.a.	V.g.	Ad.	Ad..	Ad.	Df.	N.a.	N.a.	V.g.	Df.
Krause et al. 2016	N.a.	V.g.	Ad.	Ad.	Ad.	Df.	N.a.	N.a.	V.g.	Df.
Tricot et al. 2017	N.a.	N.a.	N.a.	N.a.	N.a.	I.ad.	N.a.	N.a.	N.a.	I.ad.
Steginsky et al. 2020	N.a.	N.a.	N.a.	N.a.	N.a.	I.ad.	N.a.	N.a.	N.a.	I.ad.
Knupp et al. 2008	N.a.	N.a.	N.a.	N.a.	N.a.	I.ad.	N.a.	N.a.	N.a.	I.ad.
Martin Oliva et al. 2017	N.a.	N.a.	N.a.	N.a.	N.a.	I.ad.	N.a.	N.a.	N.a.	I.ad.
Maenohara et al. 2018	N.a.	N.a.	N.a.	N.a.	N.a.	I.ad.	N.a.	N.a.	N.a.	I.ad.
Mirza et al. 2019	N.a.	N.a.	N.a.	N.a.	N.a.	I.ad.	N.a.	N.a.	N.a.	I.ad.

N.a.: not applicable; V.g.: Very good; Ad.: Adequate; Df.: Doubtful; I.ad.: Inadequate. 1) Were patients stable in the time between the repeated measurements on the construct to be measured?; 2) Was the time interval between the repeated measurements appropriate?; 3) Were the measurement condition similar for the repeated measurements?; 4) Did the professional(s) assign the scores or determined the scores

without knowledge of scores or values of other repeated measurement(s) in the same patients?; 5) Did the professional(s) assign the scores or determined the values without knowledge of the scores or values of other repeated measurement(s) in the same patients?; 6) Were there any other important flaws in the design or statistical methods of the study?; 7) For continuous scores: was an Intraclass Correlation Coefficient (ICC) calculated?; 8) For ordinal scores: was a (weighted) Kappa calculated?; 9) For dichotomous/nominal scores: was Kappa calculated for each category against the other categories combined?

The primary aims of the studies included in this review were diverse. The primary aim of 16 included studies, was to investigate the effect of bone healing stimulating therapies, such as growth factors, bone grafts or low intensity pulsed ultrasound [10, 16-19, 23-25, 28, 29, 31, 32, 37, 39-41]. Eight studies investigated different types of osteosynthesis material or operative techniques [20-22, 27, 30, 33, 35, 44]. In four studies the primary aim was to correlate fusion ratio to clinical outcomes [13-15, 18]. In one study, the primary study aim was to evaluate the interrater reliability of consolidation assessment from CT [43]. Other studies evaluated long-term results of arthrodesis [34], risk factors for nonunion [38], bilateral versus unilateral ankle arthrodesis [36], radiographic versus CT evaluation of healing after arthrodesis [9], or fusion rate after arthrodesis [42]. See table 3 for more detailed aims of the studies and a complete overview of the general study characteristics.

Methods of consolidation assessment

Figure 2 shows an overview of methods used for consolidation assessment. Four studies did not describe the methodology used to assess consolidation [34-37]. Thirteen studies calculated fusion ratio from CT [9, 13-17, 19-21, 26, 42, 44]. The methodology of calculating fusion ratio in foot and ankle arthrodesis was first described in 2006, by the studies of Jones et al. (2006) and Coughlin et al. (2006), which are studies from the same research group [9, 10]. Fusion ratio was calculated by measuring the width of the joint surface on each CT slice, and the widths of the fused segments on each CT slice. The widths of the joint surfaces of each slice were summed, as were the widths of the fused segments on each slice. Fusion ratio was calculated with the next formula: fusion ratio = (sum of widths of fused segments / sum of widths of joint surfaces) * 100'.

The studies of Jones et al. (2006) and Coughlin et al. (2006) measured the widths of fused parts and arthrodesis with a handheld digital micrometer [9, 10]. Slices that were used for calculating fusion ratio were the first to last slice that showed parallel arthrodesis surfaces subjected to graft incorporation [9]. Only apposed portions of joints were included in the measurements, herewith excluding parts with subluxation or offset of joint surfaces and joint surfaces that were curving away from each other [9]. Also, hardware and beam hardening artifacts were excluded from measurements [9].

Table 3. Study characteristics

	Number of patients	Patient population
Jones et al. 2006	13	Hindfoot nonunions treated with revision arthrodesis
Dorsey et al. 2009	29	Ankle or subtalar joint with persistent/ recurrent pain
Thaunat et al. 2012	14	Posterior arthroscopic subtalar arthrodesis
Coulomb et al. 2019	22	Arthroscopic subtalar arthrodesis
Fourman et al. 2014	82	Ilizarov method in complex subtalar ankle arthrodesis
Mehlhorn et al. 2020	40	Medial column stabilization of instable Charcot foot
Coughlin et al. 2006	15	Hindfoot arthrodesis
Coughlin et al. 2008	15	Subtalar arthrodesis
Glazebrook et al. 2013a	275	Isolated hindfoot or ankle fusions
Usuelli et al. 2016	25	Total ankle replacement with subtalar fusion
Myerson et al. 2019	140	Subtalar arthrodesis
Ford et al. 2019	33	Tibiotalocalcaneal arthrodesis
Bejarano-Pineda 2020	7	Tibiocalcaneal hindfoot arthrodesis
Wanivenhaus et al. 2017	39	First Metatarsophalangeal joint arthrodesis with compressive locking plate
Jones et al. 2015	103	Hindfoot, midfoot or ankle arthrodesis with CBA
Daniels 2010	60	Hindfoot or midfoot arthrodesis
DiGiovanni et al. 2011	20	Ankle or hindfoot arthrodesis
Glazebrook et al. 2013b	24	Hindfoot, midfoot and ankle arthrodesis
Dekker et al. 2018	15	Tibia, ankle or hindfoot reconstructive procedure
DiGiovanni et al. 2013	434	Hindfoot or ankle arthrodesis
Daniels et al. 2019	106	Ankle or hindfoot arthrodesis
Daniels et al. 2015	75	Ankle and hindfoot fusions
Kodama et al. 2016	27	Ankle arthrodesis
Aubret et al. 2018	11	Revision of total ankle replacement by arthrodesis
Bibbo 2009	69	Ankle or hindfoot fusions in patients at high risk for nonunion

Study aim	Slice thickness (mm)	Planes used for assessment
Effect of low-intensity ultrasound in patients with nonunion treated with revision arthrodesis	1	Orthogonal to joint (ax., sag., or cor.)
Correlate joint stability with bone fusion	2	Sag.
Correlate fusion ratio and functional results	2	Sag.
Correlate between fusion ratio and functional results	NR	Sag.
Effect of rhBMP-2 on fusion rate	1-2	Sag.
High-profile threaded and surface modified fusion bolts compared to standard fusion bolts	NR	NR
Comparing radiographs and CT-scans for the quantitative evaluation of healing of hindfoot arthrodesis	2	Ax. & cor.
Effect of low intensity ultrasound bone stimulation after subtalar arthrodesis Evaluate healing rate and	2	Ax. & cor.
correlate fusion with good clinical outcomes	2	Sag.
Investigate fusion rate of subtalar joint with total ankle replacement	2	Sag. & cor.
Assess the safety and efficacy of adipose-derived cellular bone matrix compared to autograft	NR	Sag. & cor.
Effect of arthrodesis nail with internal pseudoelastic nitinol compression element	1.2-2	Sag. & cor.
Evaluate outcome of retrograde intramedullary nail and custom 3D printed titanium cage	NR	NR
Evaluate fusion rate in arthrodesis with dorsal fusion plate combined with plantar lag screw	NR	NR
Assess safety and effectiveness of CBA	NR	Sag. & cor.
Effect of rhPDGF in a calcium phosphate matrix	NR	NR
Compare safety and efficacy of biosynthetic bone graft substitute to autograft	2	Ax., cor. & sag.
Compare safety and effectiveness of B2A-granule to autograft	0.6-2	Ankle: Cor. Subtalar & talonavicular: sag.
Report outcomes of patient-specific 3D-printed titanium implants	NR	NR
Evaluate if RhPDGF-BB growth factor combined with osteoinductive matrix is a safe and effective alternative to autograft	2	Orthogonal to joint (ax., sag., or cor.)
Compare safety and efficiency of rhPDGF-BB/b-TCP- collagen with autograft	NR	NR
Evaluate efficacy and safety of rhPDGF-BB combined with beta-tricaclium phosphate colagen matrix versus autograft	NR	NR
Compare vascularized and non-vascularised anterior sliding tibial grafts	NR	NR
TAR revision by reconstruction-arthrodesis when using ankle spacer	NR	Ax., sag. & cor.
Effect of rhBMP-2 augmentation in high-risk fusions	NR	NR

Table 3. continued part 2

	Number of patients	Patient population
Cerrato et al. 2014	41	Ankle, subtalar or tibiotalocalcaneal fusion
DiGiovanni et al. 2016	397	Ankle or hindfoot arthrodesis and required supplemental bone graft
Krause et al. 2016	370	Hindfoot, midfoot and ankle fusion
Tricot et al. 2017	82	Hindfoot or ankle arthrodesis
Steginsky et al. 2020	32	Isolated subtalar ankle arthrodesis at risk of nonunion
Knupp et al. 2008	28	Rheumatoid arthritis with triple arthrodesis
Martin Oliva et al. 2017	19	Arthroscopic subtalar arthrodesis
Maenohara et al. 2018	20	Bilateral or unilateral ankle arthrodesis
Mirza et al. 2019	18	Hindfoot, midfoot, forefoot and ankle arthrodesis

Pro.: Prospective cohort; Retro.: Retrospective cohort; RCT: Randomized controlled trial; Sag.: Sagital; Ax.: Axial; Cor.: Coronal

Next to the studies of Jones et al. and Coughlin et al., 11 studies calculated fusion ratio to assess consolidation. They referred either to the study of Jones et al. [13-16, 44], or to the study of Coughlin et al. [17-21, 42]. Three of these studies used the calculated fusion ratio as primary outcome measure for consolidation [13, 14, 16].

After calculating fusion ratio, three studies categorized fusion ratios into arbitrary groups. For example, two studies classified an arthrodesis as nonunion if the fusion ratio was lower than 33%, fusion ratios between 34% to 66% were classified as partial fused, and fusion ratio of 67% and higher were classified as complete fusion [10, 15]. Mehlhorn et al (2020). categorized fusion ratios in three groups, less than 25% fusion, more than 25% and less than 50% fusion, and more than 50% fusion [44].

The remaining seven studies who calculated a fusion ratio, applied a fusion threshold to decide whether an arthrodesis was fused or not. Six studies used a fusion threshold of 50% [9, 17, 18, 20, 21, 42], one study used a fusion threshold of 45% [19].

Wanivenhaus et al. (2017) had an unique method for consolidation assessment. They divided the joint space into nine subareas, by dividing the joint space into three horizontal planes (dorsal, central and plantar) and three vertical planes (medial, central and lateral) [22]. For each subarea the bony bridging was rated and the total amount of consolidation was manually calculated. Thereafter, consolidation was categorized into no fusion (<20%), partial fusion (20%-90%) or total fusion (>90%) [22].

Study aim	Slice thickness (mm)	Planes used for assessment
Evaluate interrater reliability of fusion from CT	2	Sag. & cor.
Determine adequate graft material for fusion	2	Orthogonal to joint (ax., sag. or cor.)
Evaluate impact of nonunion on clinical outcomes and identify potential risk factors for nonunion	NR	NR
Evaluate if allograft-DBM-BMA is as effective as autograft –DBM treatment	NR	NR
Evaluate anterior ankle arthrodesis with rigid plate fixation	NR	NR
Evaluate long-term results of triple arthrodesis	NR	NR
Report outcome arthrodesis using 2 posterior portals	NR	NR
Compare outcomes of bilateral and unilateral ankle arthrodesis	NR	NR
Investigate the efficacy of LIPUS in delayed union and nonunion following foot and ankle arthrodesis	NR	NR

Sixteen studies had a more subjective approach to assess consolidation, without assessment of an objective measure. Ten studies subjectively assessed consolidation from CT by categorizing the amount of consolidation into groups [23-30, 40, 41]. Six studies subjectively assessed consolidation with a fusion threshold. Five studies used a fusion threshold of 50% [31-33, 39, 43], and one a fusion threshold of 25% [38].

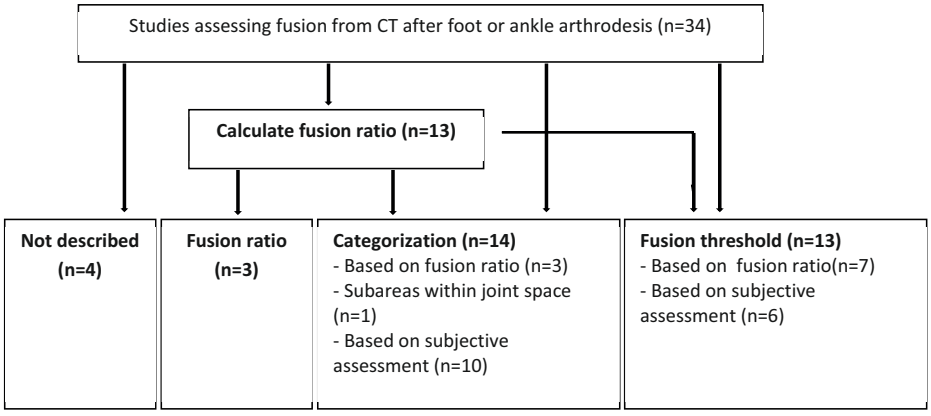


Figure 2. Methods of fusion assessment

See table 4 for an overview of all methods used for consolidation, including fusion thresholds, consolidation categories and who performed the assessment of consolidation.

Table 4. Methods of consolidation assessment

	Method of consolidation assessment	Fusion threshold	Consolidation categories	Assessor	Reference
Jones et al. 2006	Fusion ratio		≤33%: Nonunion 34%-66%: Partial fused ≥76%: Complete fusion	NR	NR
Dorsey et al. 2009	Fusion ratio			Two blinded radiologists	Jones et al. 2006
Thaunat et al. 2012	Fusion ratio			NR	Dorsey et al. 2009
Coulomb et al. 2019	Fusion ratio		≤33%: Nonunion 33%-67%: Partial fusion ≥67%: Complete	NR	Dorsey et al. 2009
Fourman et al. 2014	Fusion ratio			Blinded radiologist	Jones et al. 2006
Mehlhorn et al. 2020	Fusion ratio		<25% 25%-50% >50%	Trauma & orthopedic surgeon	Jones et al. 2006
Coughlin et al. 2006	Fusion ratio	50%		Independent radiologist	NR
Coughlin et al. 2008	Fusion ratio	50%		Independent radiologist	Coughlin et al. 2006
Glazebrook et al. 2013a	Fusion ratio	50%		Independent reviewer	Coughlin et al. 2006
Usuelli et al. 2016	Fusion ratio	50%		3 orthopedic surgeons	Coughlin et al. 2006
Myerson et al. 2019	Fusion ratio	45%		Examiner	Coughlin et al. 2006
Ford et al. 2019	Fusion ratio	50%		Two authors	Coughlin et al. 2006
Bejarano-Pineda et al. 2020	Fusion ratio	50%		Blinded foot & ankle fellow	Coughlin et al. 2006
Wanivenhaus et al. 2017	Joint space in 9 subareas, manual calculation of fusion		<20%: No fusion 20%-90%: Partial fusion >90%: Total fusion.	One observer	NR
Jones et al. 2015	Subjective assessment		<33%: nonunion 34%-66%: partial fused >76%: complete fusion	NR	NR
Daniels et al. 2010	Subjective assessment		Absent, minimal, moderate or complete	NR	NR

Table 4. continued part 2

	Method of consolidation assessment	Fusion threshold	Consolidation categories	Assessor	Reference
DiGiovanni et al 2011	Subjective assessment		0%-25%: Absent 26-50%: Minimal 51-75%: Moderate 76-100%: Complete	Independent radiologist	NR
Glazebrook et al. 2013b	Subjective assessment		0%-24%: Absent 25%-49%: Minimal 50%-74%: Moderate 75%-100%: Complete	Independent radiologist	NR
Dekker et al. 2018	Subjective assessment		Bridging by 3 of 4 cortices	Independent radiologist & senior author	Glazebrook et al. 2013b
DiGiovanni et al. 2013	Subjective assessment		0%-24%: Absent 25%-49%: Minimal 50%-74%: Moderate 75%-100%: Complete	Blinded radiologist	NR
Daniels et al. 2019	Subjective assessment		0%-24% 25%-49% 50%-74% 75%-100%	Blinded radiologist	DiGiovanni et al. 2013
Daniels et al. 2015	Subjective assessment		0%-24% 25%-49% 50%-74% 75%-100%	Blinded radiologist	Easley et al. 2000
Kodama et al. 2016	Subjective assessment		Joint line disappeared with clear trabeculation Complete Some joint line is present with partial trabeculation Moderate Entire joint space present Nonunion	NR	NR

Table 4. continued part 3

	Method of consolidation assessment	Fusion threshold	Consolidation categories	Assessor	Reference
Aubret et al. 2018	Subjective assessment		<30%: nonunion 30%-70%: doubtful fusion > 70%: definitive fusion	NR	NR
Bibbo et al. 2009	Subjective assessment	50%		NR	NR
Cerrato et al. 2014	Subjective assessment	50%		4 orthopedic foot and ankle surgeons	Coughlin et al. 2006
DiGiovanni et al. 2016	Subjective assessment	≥50%		Blinded radiologist	NR
Krause et al. 2016	Subjective assessment	≥25%		Independent radiologist	Glazebrook et al. 2013b
Tricot et al. 2017	Subjective assessment	50%		NR	NR
Steginsky et al. 2020	Subjective assessment	50%		Senior author	NR
Knupp et al. 2008	NR	NR	NR	Independent radiologist	NR
Martin Oliva et al. 2017	NR	NR	NR	NR	NR
Maenohara et al. 2018	NR	NR	NR	NR	NR
Mizra et al. 2019	NR	NR	NR	NR	NR

Reliability measures

In the study of Usuelli et al. (2016), three orthopedic surgeons were asked to objectively measure and calculate subtalar fusion ratio. Thereafter, a fusion threshold of 50% was applied. This study reported excellent interrater reliability for the subtalar joint (kappa=.91, see table 5) [42].

The study of Mehlhorn et al. (2020) included patients with an instable Charcot foot who were stabilized by fixating the first metatarsal joint, talonavicular joint and calcaneocuboid joint. Two observers (experienced trauma surgeon and orthopedic surgeon) calculated the fusion ratio for each joint and then categorized consolidation into one of three groups (<25% fusion, 25-50% fusion and >50% fusion). Based on the consolidation group, joints received points for consolidation. Zero points for less than 25% fusion, 1 point for 25-50% fusion and 2 points for > 50% fusion. As three joints were scored per patient, final consolidation scores ranged between 0 and 6 points. With this scoring system, the study showed substantial level of agreement between raters (kappa=.72, see table 5) [44].

The study of Cerrato et al. (2014) aimed to evaluate the reproducibility scores for consolidation assessment after subtalar, ankle or tibiotalocalcaneal arthrodesis. Four orthopedic surgeons were asked to subjectively judge the extent of consolidation on sagittal and coronal CT cuts. Surgeons judged whether there was more than 50% consolidation across the arthrodesis site. Kappa scores were separately reported per joint but were at least substantial (kappa>.61) [43]. See table 5 for all Kappa scores and 95% confidence intervals. The studies of DiGiovanni et al. (2016) and Krause et al. (2016) reported intrarater reliability for the same methodology. In both studies a radiologist subjectively assessed whether a joint was fused, and reassessed the CT scans after at least three months. DiGionvanni et al. (2016) applied a fusion threshold of 50%, and Krause et al. (2008) of 25%. Kappa value for the intrarater reliability was excellent in both studies (DiGiovanni et al. (2016): k=0.87; Krause et al. (2016): k=0.87, see table 5) [38, 39].

The study of DiGiovanni et al. (2013) reported intrarater reliability by having the same blinded radiologist re-examine a subset of CT scans with at least three months in between. The study was performed in patients with ankle or hindfoot arthrodesis. In this study, the radiologist subjectively categorized consolidation into one of four groups (0%-24% fusion, 25%-49% fusion, 50%-74% fusion, 75-100% fusion). Kappa value for intrarater reliability was substantial (k=0.67, see table 5) [40].

None of the other studies reported reliability scores for their methodology.

Validity measures

Five studies correlated CT assessed consolidation to the clinical outcome. Glazebrook et al. (2013b) did this in patients with isolated subtalar joint fusion. Clinical outcome was assessed with questionnaires, and consolidation was subjectively categorized as absent (0%-24%), minimal (25%-49%), moderate (50%-74%) or complete (75%-100%). The study showed that for considering a hindfoot or ankle fusion clinical successful at least minimal fusion (25%-49%) is necessary [26].

Table 5. Outcomes of reliability measures

Study	Reliability	Method of consolidation assessment	Joint of arthrodesis	Kappa (95% CI)
Usuelli et al. 2016	Interrater	Fusion ratio with 50% fusion threshold	Subtalar	0.91 (0.73-1.00)
Mehlhorn et al. 2020	Interrater	Fusion ratio with categorization	First tarsometatarsal, talonavicular & calcaneocboid	0.72 (NR)
Cerrato et al. 2014	Interrater	Subjective assessment with 50% fusion threshold	Isolated ankle TCC ankle Isolated subtalar TCC subtalar	0.79 (0.58-1.00) 0.73 (0.30-1.00 0.93 (0.74-1.00) 1.00 (0.80-1.00)
DiGiovanni et al. 2016	Intrarater	Subjective assessment with 50% fusion threshold	Foot and ankle joint	0.87 (NR)
Krause et al. 2016	Intrarater	Subjective assessment with 25% fusion threshold	Hindfoot or ankle joints	0.87 (NR)
DiGiovanni et al. 2013	Intrarater	Subjective categorization	Ankle, talonavicular, subtalar or calcaneocuboid	0.67 (0.46-0.87)

TCC: Tibiotalocalcaneal; NR: not reported

Three studies investigated the correlations between fusion thresholds and clinical outcomes. Dorsey et al. (2009) correlated fusion ratio to stability in patients with persistent or recurrent pain after ankle or subtalar joint arthrodesis [13]. Stability was assessed by physical examination, operative reports, and effect of diagnostic injections. Based on this information, a joint was judged to be stable or unstable. Twelve of 42 operated joints were judged as unstable, and 30 as stable. Unstable joints had fusion ratios ranging between 0% to 32.8%. Stable joints had fusion ratios ranging from 33.2% to 100%. Dorsey et al. (2009) set the optimal cutoff level to 33% as this threshold yielded 100% sensitivity and specificity [13].

Thaunat et al. (2012) aimed to validate the 33% fusion ratio threshold found by Dorsey et al. (2009) in patients with arthroscopic subtalar arthrodesis [14]. At six months, three out of 14 fusions had poor clinical outcome and needed revision surgery. In all three cases, fusion ratio was less than 33% fusion. All other cases had fusion ratios higher than 33% and had clinically stable fusions. Based on this descriptive study, Thaunat et al. (2012) concluded that a 33% fusion ratio threshold can indeed discriminate between clinical stable and unstable fusions. Furthermore, Fourman et al. (2014) reported in their study, that all patients who had a fusion ratio of at least 30% at three months, achieved successful consolidation without further interventions [16]. Coulumb et al. (2019) also aimed to correlate functional results to fusion ratio, but this study did not find any correlations between functional results and fusion ratio [15].

Data synthesis

Overall, the most applied method for consolidation assessment after arthrodesis in foot and ankle joints is subjective categorization (n=10), followed by calculating fusion ratio and applying a fusion threshold (n=7) and subjective assessment with fusion threshold (n=6). Less common methods were calculating fusion ratio (n=3), categorization based on fusion ratio (n=3) and creating subareas within the joint space (n=1). Four of the included studies the methodology for consolidation assessment was not described.

For some of these methods, reliability was assessed. Calculating fusion ratio and applying a fusion threshold showed excellent agreement between raters [42]. Calculating fusion ratio, categorization and addressing points for three joints in the foot to assess overall consolidation showed substantial agreement between raters [44]. Subjective assessment with fusion threshold showed substantial to excellent agreement between raters [43] and excellent agreement within the same rater [38, 39]. Subjective categorization showed substantial reliability within the same rater [40].

For studies with subjective categorization, one study reported that at least minimal (25%-49%) fusion is necessary for good clinical outcome [26]. For fusion threshold, three studies reported that fusion ratio's ranging from 30% to 33% are associated with clinically stable joints [13, 14, 16]. One study reported no correlations between fusion ratio and functional results [15].

Discussion

This review showed that there is a lot of heterogeneity in the assessment of consolidation from CT in foot and ankle arthrodesis. Although consolidation was the most important outcome measure in 88% of the included studies, 80% of the studies were at high risk of bias due to a poor methodology of consolidation assessment. Only seven of 34 studies reported, or referred to studies with, outcomes related to the reliability of the method that was used.

For the three most frequently used methods, reliability was assessed by at least one study. The most popular method for consolidation assessment is subjective categorizing of consolidation. Although the reliability was substantial for this method, the subjectivity may decrease the reliability of this method [40]. Another popular method was to calculate the fusion ratio and then apply a fusion threshold. This more objective method resulted in excellent agreement between raters [42]. Subjective fusion rate assessment with a fusion threshold also resulted in good reliability scores. Other methods for consolidation assessment were used less frequently and were also not assessed for reliability.

For future studies we would strongly recommend to calculate fusion ratio for the assessment of consolidation after foot or ankle arthrodesis, as this seems to be the most objective method. The methodology of measuring and calculating fusion ratio from CT was first described by Singh et al. (2005) in patients with scaphoid fractures [45]. In 2013, the reliability of fusion ratio as a continuous outcome measure was assessed in patient with scaphoid fractures, which was excellent between two raters [46]. For consolidation assessment in patients with foot or ankle arthrodesis, it would also be very interesting to know if fusion ratio could reliably represent consolidation on a continuous scale ranging from 0-100% fusion. However, the reliability of fusion ratio has not been assessed in this patient population yet. Also, currently fusion ratio is assessed with varying slice thicknesses and viewing planes (sagittal, coronal and axial). Based on the current studies, it is unclear which slice thicknesses and which plane(s) should be used. Table 3 showed that most studies who reported slices thickness used slice thickness' of 2 mm. Future studies may investigate whether thicker slices can be used, without losing accuracy. Which planes can best be used is dependent on the joint that is assessed. For the subtalar joint, most studies used only the sagittal plane. However, some studies used the sagittal and coronal planes. We think that consolidation is most accurately assessed if we look at several planes, as is common in clinical practice. We therefore advise to use the two planes that are orthogonal to the joint that is assessed, and to take the average fusion ratio of the two planes. For example, for the subtalar joint this would be the average of the sagittal and coronal planes.

Only three studies used the fusion ratio as their primary outcome measure for consolidation [13, 14, 16]. Most studies categorized the amount of consolidation into consolidation groups, or set a fusion threshold to distinguish between fused and non-fused joints. Choosing appropriate boundaries for consolidation groups or fusion thresholds is challenging. Most studies in this review used a fusion threshold of 50%, which was rather randomly chosen [9, 17, 18, 20, 21, 31-33, 39, 42, 43]. Studies who correlated consolidation status to clinical outcomes showed that good clinical outcomes can be expected from joints that have been fused for more than 30% [13, 14, 16, 26]. It seems that many studies use fusion thresholds that are quite high, and therefore, the percentage of fused joints may be underestimated in those studies. However, the four studies that have been done on validity were all quite small, with a total of 149 participants in all four studies together. Also, all the studies were performed in patients with subtalar or ankle arthrodesis, so it is unclear if a 30% fusion threshold is also applicable in the smaller joints of the foot. It would be interesting to validate the 30% fusion threshold in a future study, with large numbers and good methodological set-up. However, with the limited amount of available evidence, we would advise to use a 30% fusion threshold to distinguish fused from non-fused joints. Also, based on this fusion threshold, it may be most appropriate to use consolidation categories where ≤33% represent nonunion, 34-66% partial fusion, and ≥76% complete fusion.

This review encountered some limitations. Firstly, eligible studies were selected based on mentioning of radiological or CT follow-up in the abstract. Herewith, we might have missed studies who did use CT for consolidation assessment but did not mention those terms in their abstract. Secondly, the included studies provided limited information about who performed the assessment of consolidation and about the assessors' experience. For future studies it would be of interest to investigate the reliability of fusion assessment between more experienced assessors, like musculoskeletal radiologists, and less experienced assessors. Lastly, this systematic review advises on radiological consolidation assessment in foot and ankle joints. However, most of the included studies were about arthrodesis of the ankle, hindfoot or midfoot joints. Only two studied assessed consolidation in the metatarsophalangeal joint in the forefoot. Mizra et al. (2019) did not report how consolidation was assessed, and Wanivenhaus et al. (2017) divided the joint space in nine subareas to assess consolidation. Whether calculating fusion ratio is a reliable and valid method in the small joints of the forefoot is therefore unclear and should be assessed in future studies.

Overall, clinimetric evidence for the best method of consolidation assessment from CT in foot and ankle arthrodesis is limited. However, based on the current literature, we would like to make some recommendations for future studies. We advise to measure

and calculate fusion ratio as an outcome measure after foot and ankle arthrodesis [9, 10]. In this method, fusion ratio is calculated with the formula: 'sum of widths of fused segments / sum of widths of joint surfaces * 100'. Preferably, two independent assessors should assess fusion thresholds. CT slices should be set at a maximum slice thickness of 2 mm. The assessment should be done based on the two orthogonal planes, and fusion ratio should be assessed by taking the average of those two planes. Based on the studies presented in this systematic review, we would advise to use 33% as a fusion threshold to discriminate between non-fused and fused joints, or categorize consolidation outcomes in three groups (≤33% represent nonunion, 34-66% partial fusion, and ≥76% complete fusion). Future clinimetric studies should focus on assessing the reliability of fusion ratio as a continuous outcome measure and validating the 33% fusion threshold.

References

[1] Yasui Y, Hannon CP, Seow D, Kennedy JG. Ankle arthrodesis: A systematic approach and review of the literature. World J Orthop. 2016;7:700-8.

[2] Rabinovich RV, Haleem AM, Rozbruch SR. Complex ankle arthrodesis: Review of the literature. World J Orthop. 2015;6:602-13.

[3] Corrales LA, Morshed S, Bhandari M, Miclau T, III. Variability in the Assessment of Fracture-Healing in Orthopaedic Trauma Studies. JBJS. 2008;90.

[4] Whelan DB, Bhandari M, Stephen D, Kreder H, McKee MD, Zdero R, et al. Development of the Radiographic Union Score for Tibial Fractures for the Assessment of Tibial Fracture Healing After Intramedullary Fixation. Journal of Trauma and Acute Care Surgery. 2010;68.

[5] Cooke ME, Hussein AI, Lybrand KE, Wulff A, Simmons E, Choi JH, et al. Correlation between RUST assessments of fracture healing to structural and biomechanical properties. J Orthop Res. 2018;36:945-53.

[6] Litrenta J, Tornetta P, 3rd, Mehta S, Jones C, O'Toole RV, Bhandari M, et al. Determination of Radiographic Healing: An Assessment of Consistency Using RUST and Modified RUST in Metadiaphyseal Fractures. J Orthop Trauma. 2015;29:516-20.

[7] Misir A, Uzun E, Kizkapan TB, Yildiz KI, Onder M, Ozcamdalli M. Reliability of RUST and Modified RUST Scores for the Evaluation of Union in Humeral Shaft Fractures Treated with Different Techniques. Indian J Orthop. 2020;54:121-6.

[8] Leow JM, Clement ND, Simpson A. Application of the Radiographic Union Scale for Tibial fractures (RUST): Assessment of healing rate and time of tibial fractures managed with intramedullary nailing. Orthop Traumatol Surg Res. 2020;106:89-93.

[9] Coughlin MJ, Grimes JS, Traughber PD, Jones CP. Comparison of radiographs and CT scans in the prospective evaluation of the fusion of hindfoot arthrodesis. Foot Ankle Int. 2006;27:780-7.

[10] Jones CP, Coughlin MJ, Shurnas PS. Prospective CT scan evaluation of hindfoot nonunions treated with revision surgery and low-intensity ultrasound stimulation. Foot Ankle Int. 2006;27:229-35.

[11] Mokkink LB, Boers M, van der Vleuten CPM, Bouter LM, Alonso J, Patrick DL, et al. COSMIN Risk of Bias tool to assess the quality of studies on reliability or measurement error of outcome measurement instruments: a Delphi study. BMC Medical Research Methodology. 2020;20:293.

[12] Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33:159-74.

[13] Dorsey ML, Liu PT, Roberts CC, Kile TA. Correlation of arthrodesis stability with degree of joint fusion on MDCT. Am J Roentgenol. 2009;192:496-9.

[14] Thaunat M, Bajard X, Boisrenoult P, Beaufils P, Oger P. Computer Tomography assessment of the fusion rate after posterior arthroscopic subtalar arthrodesis. Int Orthop. 2012;36:1005-10.

[15] Coulomb R, Hsayri E, Nougarede B, Marchand P, Mares O, Kouyoumdjian P, et al. Do clinical results of arthroscopic subtalar arthrodesis correlate with CT fusion ratio? Orthop Traumatol : Surg Res. 2019;105:1125-9.

[16] Fourman MS, Borst EW, Bogner E, Rozbruch SR, Fragomen AT. Recombinant human BMP-2 increases the incidence and rate of healing in complex ankle arthrodesis. Clin Orthop Relat Res. 2014;472:732-9.

[17] Coughlin MJ, Smith BW, Traughber P. The evaluation of the healing rate of subtalar arthrodeses, part 2: the effect of low-intensity ultrasound stimulation. Foot Ankle Int. 2008;29:970-7.

[18] Glazebrook M, Younger A, Wing K, Lalonde KA. A prospective pilot study of b2a-coated ceramic Granules (Amplex) compared to autograft for ankle and hindfoot arthrodesis. Foot Ankle Int. 2013A;34:1055-63.

[19] Myerson CL, Myerson MS, Coetzee JC, Stone McGaver R, Giveans MR. Subtalar Arthrodesis with Use of Adipose-Derived Cellular Bone Matrix Compared with Autologous Bone Graft: A Multicenter, Randomized Controlled Trial. J Bone Joint Surg Am. 2019;101:1904-11.

[20] Ford SE, Kwon JY, Ellington JK. Tibiotalocalcaneal Arthrodesis Utilizing a Titanium Intramedullary Nail With an Internal Pseudoelastic Nitinol Compression Element: A Retrospective Case Series of 33 Patients. J Foot Ankle Surg. 2019;58:266-72.

[21] Bejarano-Pineda L, Sharma A, Adams SB, Parekh SG. Three-Dimensional Printed Cage in Patients With Tibiotalocalcaneal Arthrodesis Using a Retrograde Intramedullary Nail: Early Outcomes. Foot Ankle Spec. 2020:1938640020920947.

[22] Wanivenhaus F, Espinosa N, Tscholl PM, Krause F, Wirth SH. Quality of Early Union After First Metatarsophalangeal Joint Arthrodesis. J Foot Ankle Surg. 2017;56:50-3.

[23] Jones CP, Loveland J, Atkinson BL, Ryaby JT, Linovitz RJ, Nunley JA. Prospective, Multicenter Evaluation of Allogeneic Bone Matrix Containing Viable Osteogenic Cells in Foot and/or Ankle Arthrodesis. Foot Ankle Int. 2015;36:1129-37.

[24] Daniels T, DiGiovanni C, Lau JTC, Wing K, Alastair Y. Prospective clinical pilot trial in a single cohort group of rhPDGF in foot arthrodeses. Foot Ankle Int. 2010;31:473-9.

[25] DiGiovanni CW, Baumhauer J, Lin SS, Berberian WS, Flemister AS, Enna MJ, et al. Prospective, randomized, multi-center feasibility trial of rhPDGF-BB versus autologous bone graft in a foot and ankle fusion model. Foot Ankle Int. 2011;32:344-54.

[26] Glazebrook M, Beasley W, Daniels T, Evangelista PT, Donahue R, Younger A, et al. Establishing the relationship between clinical outcome and extent of osseous bridging between computed tomography assessment in isolated hindfoot and ankle fusions. Foot Ankle Int. 2013B;34:1612-8.

[27] Dekker TJ, Steele JR, Federer AE, Hamid KS, Adams SB, Jr. Use of Patient-Specific 3D-Printed Titanium Implants for Complex Foot and Ankle Limb Salvage, Deformity Correction, and Arthrodesis Procedures. Foot Ankle Int. 2018;39:916-21.

[28] Daniels TR, Younger ASE, Penner MJ, Wing KJ, Le ILD, Russell IS, et al. Prospective randomized controlled trial of hindfoot and ankle fusions treated with rhPDGF-BB in combination with a β-TCP-collagen matrix. Foot Ankle Int. 2015;36:739-48.

[29] Kodama N, Takemura Y, Shioji S, Imai S. Arthrodesis of the ankle using an anterior sliding tibial graft for osteoarthritis secondary to osteonecrosis of the talus. Bone Jt J. 2016;98B:359-64.

[30] Aubret S, Merlini L, Fessy M, Besse JL. Poor outcomes of fusion with Trabecular Metal implants after failed total ankle replacement: Early results in 11 patients. Orthop Traumatol : Surg Res. 2018;104:231-7.

[31] Bibbo C, Patel DV, Haskell MD. Recombinant bone morphogenetic protein-2 (rhBMP-2) in high-risk ankle and hindfoot fusions. Foot Ankle Int. 2009;30:597-603.

[32] Tricot M, Deleu PA, Detrembleur C, Leemrijse T. Clinical assessment of 115 cases of hindfoot fusion with two different types of graft: Allograft + DBM + bone marrow aspirate versus autograft + DBM. Orthop Traumatol : Surg Res. 2017;103:697-702.

[33] Steginsky BD, Suhling ML, Vora AM. Ankle Arthrodesis With Anterior Plate Fixation in Patients at High Risk for Nonunion. Foot Ankle Spec. 2020;13:211-8.

[34] Knupp M, Skoog A, Törnkvist H, Ponzer S. Triple arthrodesis in rheumatoid arthritis. Foot Ankle Int. 2008;29:293-7.

[35] Martín Oliva X, Falcão P, Fernandes Cerqueira R, Rodrigues-Pinto R. Posterior Arthroscopic Subtalar Arthrodesis: Clinical and Radiologic Review of 19 Cases. J Foot Ankle Surg. 2017;56:543-6.

[36] Maenohara Y, Taniguchi A, Tomiwa K, Tsuboyama D, Kurokawa H, Kumai T, et al. Outcomes of Bilateral vs Unilateral Ankle Arthrodesis. Foot Ankle Int. 2018;39:530-4.

[37] Mirza YH, Teoh KH, Golding D, Wong JF, Nathdwarawala Y. Is there a role for low intensity pulsed ultrasound (LIPUS) in delayed or nonunion following arthrodesis in foot and ankle surgery? Foot Ankle Surg. 2019;25:842-8.

[38] Krause F, Younger AS, Baumhauer JF, Daniels TR, Glazebrook M, Evangelista PT, et al. Clinical Outcomes of Nonunions of Hindfoot and Ankle Fusions. J Bone Joint Surg Am. 2016;98:2006-16.

[39] DiGiovanni CW, Lin SS, Daniels TR, Glazebrook M, Evangelista P, Donahue R, et al. The importance of sufficient graft material in achieving foot or ankle fusion. J Bone Jt Surg Am Vol. 2016;98:1260-7.

[40] DiGiovanni CW, Lin SS, Baumhauer JF, Daniels T, Younger A, Glazebrook M, et al. Recombinant human platelet-derived growth factor-BB and beta-tricalcium phosphate (rhPDGF-BB/β-TCP): An alternative to autogenous bone graft. J Bone Jt Surg Ser A. 2013;95:1184-92.

[41] Daniels TR, Anderson J, Swords MP, Maislin G, Donahue R, Pinsker E, et al. Recombinant Human Platelet–Derived Growth Factor BB in Combination With a Beta-Tricalcium Phosphate (rhPDGF-BB/β-TCP)-Collagen Matrix as an Alternative to Autograft. Foot Ankle Int. 2019;40:1068-78.

[42] Usuelli FG, Maccario C, Manzi L, Gross CE. Clinical Outcome and Fusion Rate Following Simultaneous Subtalar Fusion and Total Ankle Arthroplasty. Foot Ankle Int. 2016;37:696-702.

[43] Cerrato RA, Aiyer AA, Campbell J. Reproducibility of computed tomography to evaluate ankle and hindfoot fusions. Foot & ankle …. 2014.

[44] Mehlhorn AT, Ugland KI, Hörterer H, Gottschalk O, Südkamp N, Walther M. A high-profile thread with grit-blasted and acid-etched surface reduces loosening of medial column fusion bolt in instable Charcot foot. Foot Ankle Surg. 2020;26:637-43.

[45] Singh HP, Forward D, Davis TR, Dawson JS, Oni JA, Downing ND. Partial union of acute scaphoid fractures. J Hand Surg Br. 2005;30:440-5.

[46] Grewal R, Frakash U, Osman S, McMurtry RY. A quantitative definition of scaphoid union: determining the inter-rater reliability of two techniques. Journal of Orthopaedic Surgery and Research. 2013;8:28.

Chapter 6

Extracorporeal shock wave treatment for delayed-union and nonunion fractures; a systematic review

A. Willems, O.P. van der Jagt, D.E. Meuffels
J Orthop Trauma. 2019 Feb;33(2):97-103

Abstract

Objectives: Nonunions after bone fractures are usually treated surgically with risk of infections and failure of osteosynthesis. A noninvasive alternative is extracorporeal shock wave treatment (ESWT), which potentially stimulates bone regeneration. Therefore this review investigates whether ESWT is an effective and safe treatment for delayed-unions and nonunions.
Data sources: Embase.com, Medline ovid, Cochrane, Web-of-science, Pubmed publisher and Google scholar were systematically searched.
Study selection: Inclusion criteria included studies with patients with delayed-union or nonunion treated with ESWT; inclusion of ≥10 patients; follow-up period ≥6 weeks.
Data extraction: Assessment for risk of bias was done by two authors using the Cochrane tool. Union rates and adverse events were extracted from the studies.
Data synthesis: Two RCT's and 28 non-randomized studies were included. One RCT was assessed at medium risk of bias, and reported similar union rates between ESWT-treated patients (71%) and surgery-treated patients (74%). The remaining 29 studies were at high risk of bias due to poor description of randomization (n=1), non-randomized allocation to control groups (n=2) or absence of control groups (n=26). The average union rate after ESWT in delayed-unions was 86%, in nonunions 73%, and in nonunions after surgery 81%. Only minor adverse events were reported after ESWT.
Conclusions: ESWT seems to be effective for the treatment of delayed-unions and nonunions. However, the quality of most studies is poor. Therefore, we strongly encourage conducting well-designed RCT's to prove the effectiveness of ESWT, and potentially improve the treatment of nonunions as ESWT might be as effective as surgery but safer.
Level of Evidence: Therapeutic Level II

Introduction

Delayed-unions and nonunions are failures of bony healing after fractures, osteotomies or arthrodesis. In practice a wide variety exists in the exact definition of delayed-unions and nonunions depending on fracture site and criteria used for the assessment of bony union.[1] In this review, we define delayed-unions as fractures that do not show radiological union three months after a fracture, and nonunions as fractures that do not show radiological union six months after a fracture.

Literature shows that 3-5% of all fractures evolve into a nonunion, with highest nonunions rates reported in fractures of the scaphoid (16%), tibia (14%) and femur (14%).[2, 3] Patients with nonunions suffer from pain and decreased function, which affects a patient's daily routines and decreases their quality of life.[4, 5]

Currently most nonunions are treated with surgery, which is considered to be the "golden standard".[6] Surgical treatment options of nonunions are overall quite successful, with union rates reported between 74% to 95%.[7-10] However, complications can occur such as infection (5%), neurovascular damage (7%) or implant related problems requiring an additional surgery (5%). [7, 11, 12] Alternatively to surgery patients could be treated noninvasively, which could reduce the risk of these complications.

A noninvasive treatment for delayed-unions and nonunions is extracorporeal shock wave therapy (ESWT). ESWT is a well-known treatment for fragmentation of kidney stones, but over the last decades ESWT has been increasingly used for bone growth stimulation. In 1991 Valchanou and Michailov used ESWT for the treatment of delayed-unions and nonunions and reported bony union in 70 of 82 fractures without any complications.[13] Subsequently to these promising results, more studies have been published in which ESWT was used for delayed-union and nonunion treatment.

Bone healing after ESWT might be stimulated due to an increase in neovascularization and an upregulation of angiogenic and osteogenic growth factors.[14] Animal studies reported an increase of several growth factors after ESWT, which are important in bone regeneration (VEGF, TGF-beta 1 and BMP's).[15, 16] Also, it has been show that ESWT leads to an increased differentiation of bone marrow stem cells towards osteoprogenitor cells,[17] and thickening of the cambium layer of the periosteum by proliferation of osteoprogenitor cells.[18] Although the exact working mechanisms of ESWT is still unclear it has been hypothesized that the biological responses after ESWT are triggered by mechanotransduction, a process in which cells transform mechanical stimuli into biochemical signals.[19] During ESWT pressure waves are generated by a piezoelectric,

electromagnetic or electrohydraulic mechanism. The created pressure waves are characterized by a fast pressure rise, exposing tissue cells to shear and tensile forces. These forces might cause liberation of messengers from the extracellular matrix, which can activate genes in the cell nucleus, which induces an upregulation of growth factors.[6, 19]

In 2010, Zelle et al.[20] published a systematic review concerning the treatment of delayed-unions and nonunions with ESWT. They reported that treatment of delayed-unions and nonunions with ESWT was successful in approximately 75% of the fractures. However, this conclusion was based on ten cohort studies, which provided a poor level of evidence, and a risk of bias assessment was not performed. Presently, the clinical application of ESWT for delayed-unions and nonunions has not widely spread, although more studies have been published since the review of Zelle et al..[20] Therefore, the aim of this systematic review is to provide a comprehensive overview of the currently available literature concerning the effectiveness and safeness of ESWT in the clinical treatment of delayed-unions and nonunions.

Method

The protocol of this systematic review was prospectively registered in the International prospective register of systematic reviews (http://www.crd.york.ac.uk/prospero/; registration number CRD42016046120).

Eligibility Criteria

For this review we included studies that treated delayed-unions or nonunion with ESWT. See table 1 for a full-overview of all eligibility criteria.

Table 1. Eligibility criteria

Inclusion criteria	Exclusion criteria
• Patients with a delayed-union or nonunion who are treated with ESWT • Skeletally mature patients • An outcome measure quantifying bony union (x-ray, CT-scan) should be reported • Full text available in English, Spanish, German or Dutch • Peer reviewed study • (Randomized) controlled trials, prospective and retrospective cohort-studies	• Follow-up period < 6 weeks • Less than 10 patients

Literature search

Six databases were systematically searched on the 10th of August 2017. The databases that were searched were Embase.com, Medline ovid, Cochrane, Web-of-science, Pubmed publisher and Google scholar. The search strategy that was used for the search of Medline Ovid is presented in table 2, and was adapted for the search of the other databases. Also, reference lists of eligible articles were checked for eligible articles that were missed by our search strategy.

Table 2. Search strategy for Medline

("High-Energy Shock Waves"/ OR (((shock OR pressure*) AND wave*) OR shockwave* OR eswt OR orthotrip*)) AND (exp "Bone and Bones"/ OR exp "Bone Development"/ OR exp "Bone Remodeling"/ OR exp "Fractures, Bone"/ OR "Fracture Healing"/ OR "Bone Density"/ OR exp "Bone Diseases"/ OR (bone OR bones OR fracture* OR nonunion OR ((non OR delay*) ADJ3 (union*)) OR osteo* OR osseous OR intraosseous OR (avascular* ADJ3 necro*) OR skelet* OR pseudarthrit* OR pseudoarthrit* OR (pseud* ADJ arthrit*)).ab,ti.) NOT (exp animals/ NOT humans/)

Study selection

Articles that were found by multiple databases were deduplicated. The articles were then included or excluded based on the eligibility criteria. Articles were first screened based on title and abstract. Eligible articles were again judged based on full-text. Both selection rounds were independently performed by two reviewers (AW and OJ). After each selection round the reviewers compared their selected articles, and disagreements were discussed and resolved by consensus. A third reviewer (DM) was asked in case of an unsolved disagreement.

Risk of bias assessment

Risk of bias assessment was independently performed by two reviewers (AW and DM), using the Cochrane Risk of Bias tool for RCT's.[21] This tool contains six items, which can be scored as low, high or unclear risk of bias. The six items concern random sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessment, handling of incomplete outcome data, and selective outcome reporting. Discrepancies between the reviewers were discussed and resolved by consensus. Afterwards, studies were classified as being at low, moderate or high risk of bias. Studies were at low risk of bias if all 6 items were scored as low risk of bias. Moderate risk of bias was defined as ≥4 items scored as low risk of bias. Studies were at high risk of bias if <4 items were scored as low risk of bias.

Data extraction

A data extraction sheet was established by the reviewers (AW, OJ, DM) for accurate data extraction. Data that were extracted are general characteristics of the participants including fracture sites according to the OTA classification[22], general characteristics of the ESWT, union rates and adverse events. Data were extracted from the studies by one reviewer (AW), who also completed a full check of the extracted data after the data extraction was completed.

Primary outcome

Our primary outcome is the union rate six months after ESWT. In studies that only reported absolute numbers of bony union, union rates were calculated. If the union rate after six months was not reported, union rate was reported as has been done in the study (e.g. union rate and average healing time).
The results of the studies will be presented based on the outcome of the risk of bias assessment (low risk of bias, moderate risk of bias and high risk of bias).

Secondary outcome

Our secondary outcome is the safety of ESWT. The number of adverse events and the kind of adverse events were extracted from the studies. Adverse events were graded based on the adapted Clavien-Dindo classification.[23, 24] The Clavien-Dindo classification is a tool established in general surgery to grade the severity of complications after surgery. In this tool, adverse events are graded from 1-5, in which grade 1 indicates any deviation from normal postoperative course without the need for any additional interventions, and grade 5 is the most serious, indicating the death of a patient.

Results

Literature search

The search resulted in 2780 studies, but after deduplication 1868 studies remained for screening. The screening based on title and abstract resulted in 73 potentially eligibly studies. After reading the full texts of those studies, 30 studies were found to be eligible. Screening of the reference lists of those articles did not result in any additional studies and therefore 30 studies were included in this review (figure 1).

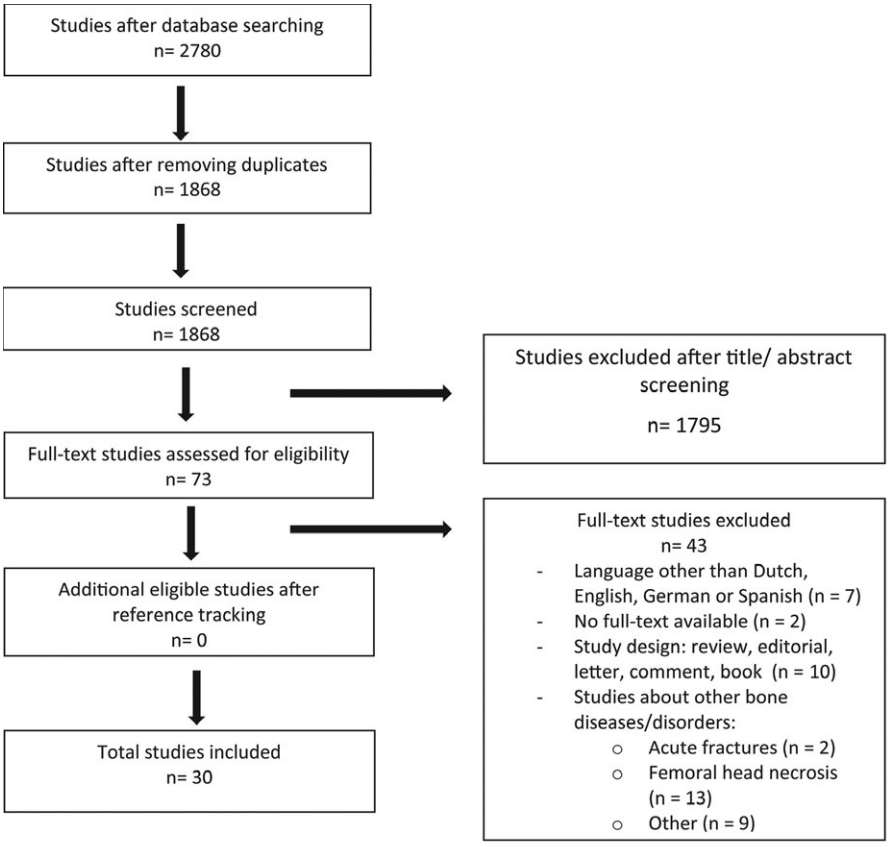


Figure 1. Flow chart of study selection

Risk of bias assessment

All 30 studies were assessed for risk of bias. After initial assessment, 173 of the 180 items were given the same score by both reviewers, and discrepancies were resolved by consensus. The results of the risk of bias assessment per study are shown in table 3. None of the studies was classified as being at low risk of bias. One study was classified as being at moderate risk of bias.[8] Twenty-nine studies were classified as being at high risk of bias.[9, 10, 13, 25-50]

Table 3. Results of risk of bias assessment of the individual studies with scores per item.

	Items of the Cochrane Risk of Bias tool for RCT's						Overall risk of bias
	1	2	3	4	5	6	
Cacchio[8] 2009	+	+	-	+	+	?	Moderate risk of bias
Zhai[49] 2016	?	?	-	?	+	?	High risk of bias
Notarnicola[10] 2010	-	-	-	-	?	?	High risk of bias
Furia[9] 2010	-	-	-	-	+	?	High risk of bias
Schaden[45] 2004	-	-	-	-	?	?	High risk of bias
Stojadinovic[39] 2011	-	-	-	-	?	?	High risk of bias
Schaden[29] 2001	-	-	-	-	?	?	High risk of bias
Everding[50] 2016	-	-	-	-	+	?	High risk of bias
Biederman[44] 2003	-	-	-	-	-	?	High risk of bias
Vulpiani[40] 2012	-	-	-	-	?	?	High risk of bias
Vogel[25] 1997a	-	-	-	-	?	?	High risk of bias
Alkhawashki[42] 2015	-	-	-	-	?	?	High risk of bias
Rodríguez de Oya[27] 2011	-	-	-	-	?	?	High risk of bias
Wang[30] 2001	-	-	-	-	-	?	High risk of bias
Xu[47] 2009	-	-	-	-	-	?	High risk of bias
Beutler[37] 1999	-	-	-	-	?	?	High risk of bias
Vogel[26] 1997b	-	-	-	-	?	?	High risk of bias
Schoellner[31]2002	-	-	-	-	?	?	High risk of bias
Rompe[28] 2001	-	-	-	-	?	?	High risk of bias
Elster[36]2010	-	-	-	-	-	?	High risk of bias
Haffner[48] 2016	-	-	-	-	-	?	High risk of bias
Wang[51]2009	-	-	-	-	?	?	High risk of bias
Kuo[43] 2015	-	-	-	-	?	?	High risk of bias
Moretti[34] 2009a	-	-	-	-	?	?	High risk of bias
Valchanou[13] 1991	-	-	-	-	?	?	High risk of bias
Bara[32] 2007	-	-	-	-	-	?	High risk of bias
Czarnowska-Cubala[31] 2013	-	-	-	-	?	?	High risk of bias
Moretti[34] 2009b	-	-	-	-	?	?	High risk of bias
Alvarez[38] 2011	-	-	-	-	-	?	High risk of bias
West[46] 2008	-	-	-	-	?	?	High risk of bias

1= Random sequence generation; 2 = Allocation concealment; 3= Blinding of participants and personnel; 4= Blinding of outcome assessment; 5= Incomplete outcome data; 6= Selective reporting; + = low risk of bias; - = high risk of bias; ? = unclear risk of bias

General characteristics of included studies

Studies at moderate risk of bias

In the RCT of Cacchio et al.[8], 126 patients with nonunions of the long bones were randomly assigned to ESWT group 1, ESWT group 2 or to the surgical treatment group. Patients in ESWT group 1 were treated with an energy flux density of 0.40 mJ/mm^2, patients in ESWT group 2 with an energy flux density of 0.70 mJ/mm^2. The general characteristics of this RCT are shown in table 4a.

Studies at high risk of bias

From the 29 studies with a high risk of bias, one study was a RCT. In this RCT, 63 patients with nonunions of the long bones were randomly assigned to ESWT combined with transplantation of human autologous bone mesenchymal stems cells (hBMSCs) or only ESWT (table 4a).[48]

Two studies at high risk of bias were retrospective non-randomized trials.[9, 10] Both studies compared patients with nonunions that were treated with ESWT, with a surgery-treated control group. The general characteristics of the patients and ESWT-treatment are shown in table 4a.

The remaining 26 studies were cohort studies without a control group, see table 4b for the general characteristics of the patients and of the ESWT-treatment. Nineteen of those studies treated nonunions in which the fracture was older than six months.[25-31, 35, 36, 38, 39, 41-44, 46, 47, 49, 50] Five of them also reported data on the treatment of delayed-unions, [29, 38, 43, 44, 49] however Everding et al. [49] was excluded from the results as they treated <10 delayed-unions. Seven studies did not report how they defined delayed-unions or nonunions.[13, 32-34, 37, 40, 45]

Of the 26 studies, ten studies collected data prospectively,[25, 26, 28, 30, 31, 36, 38, 39, 44, 47] five studies retrospectively, [34, 35, 41, 42, 49] and ten studies did not report if data were collected prospectively or retrospectively. [13, 27, 29, 32, 33, 40, 43, 45, 46, 50]

The general design of the cohort studies was that patients with delayed-unions or nonunions were treated with ESWT and were followed over time to see whether bony union did occur.

Table 4a. Characteristics of included controlled trials

		Number of fractures	Age in years (range)	Males in percentages	Fracture sites	Average months of non-union before ESWT (range)
Randomized controlled trial						
Cacchio[8] 2009	Intervention1	42	42.8 (NR)	76	Long bones of the upper and lower extremities	11.5 (NR)
	Intervention 2	42	43.1 (NR))	74		10.8 (NR)
	Control	42	42.5 (NR)	71		10.2 (NR)
Zhai[48] 2016	Intervention	32	39,6 (23 - 50)	56	Shaft of the long bones	13,4 (9 -20)
	Control	31	38,1 (20 – 49)	58		12,9 (9 –19)
Non-randomized controlled trials						
Notarnicola [10] 2010	Intervention	58	33,2 (16-65)	91	Scaphoid	14,8 (9-36)
	Control	60	33,1 (16-65)	87		15,8 (9-40)
Furia[9] 2010	Intervention	23	42,7 (17-78)	57	Proximal metaphyseal-diaphyseal part of the fifth metatarsal	10,4 (6-39)
	Control	20	40,8 (19-78)	40		6,2 (4-13)

Table 4b. General characteristics of the included cohort studies

	Number of fractures	Age in years (range)	Males in percentage	Fracture sites	Average months of non-union before ESWT (range)
Studies with delayed-unions (<6 months since fracture) and nonunions (≥ 6 months since fracture)					
Schaden[44] 2004	DU: 152 NU: 445	44 (10-90)	68	Upper + lower extremities	16,1 (NR)
Stojadinovic[38] 2011	DU: 120 NU: 229	48 (15-91)	67	Upper + lower extremities	NR
Schaden[29] 2001	DU: 35 NU: 80	43 (10-86))	64	Upper + lower extremities	NR
Everding[49] 2016	DU:9[a] NU: 33	43 (18-74)	72	Upper + lower extremities	DU: 4,3 (3,5-5) NU:17,3 (6-48)
Biederman[43] 2003	DU: 16 NU: 57	42 (NR)	53	Long bones: 58; others 12	DU: 5 (0,2-5) NU: 19 (6-74)

Percentage previously infected nonunions	Type of nonunion in percentages			Type of anesthesia	Treatment
	Hyper-thropic	Atro-phic	Oligo-tropic		
NR	71	29	0	R	ESWT (4x 4000 shocks at 0,40 mJ/mm^2)
	74	26	0	R	ESWT (4x 4000 shocks at 0,70 mJ/mm^2)
	74	26	0	G	Surgical revision with fixation (locked intramedullary nail ± autogenous bone graft or external fixator)
NR	91	90	NR	NR	Transplantation of autologous hBMSC + 4-5x ESWT (mean of 2900 shock waves at average 23 kV)
	90	10	NR	NR	4-5x ESWT (mean of 2900 shock waves at average 23 kV)
NR		NR		NR	ESWT (3x 4000 shocks at 0,05-0,12 mJ/mm^2)
				NR	Surgery according to the Matti-Russe method
NR		NR		G (n=15), R (n=6), L (n=2)	ESWT (2000-4000 shocks at 0.35 mJ/mm^2)
				NR	Intramedullary screw fixation

Percentage previously infected nonunions	Type of nonunion in percentages			Type of anesthe-sia	Number of shocks applied	Energy density in mJ/mm^2
	Hyper-thropic	Atro-phic	Oligo-trophic			
10	39	61		G, R or L	2000-4000	0,38
NR	NR	NR	NR	G or R	NR	NR
19	NR	NR	NR	G (n=60) R (n=51) L (n=4)	1000-12.000	0,25-0,40
NR	52	48	NR	NR	3000	0,36
NR	DU: 86 NU: 61	DU: 14 NU: 39	NR	G (n=46), L (n=39)	1-2x 2900 (mean)	23 kV

Table 4b. continued part 2

	Number of fractures	Age in years (range)	Males in percentage	Fracture sites	Average months of non-union before ESWT (range)
Studies with nonunions (≥ 6 months since fracture)					
Vulpiani[39] 2012	143	41 (14-81)	64	Upper + lower extremities	14,1 (6-84)
Vogel[25] 1997a	52	37 (12-81)	62	Upper + lower extremities	13 (6-51)
Alkhawashki[41] 2015	49	34 (14-70)	80	Upper + lower extremities	11,9 (6-60)
Rodríguez de Oya[27] 2001	20	42 (26-62)	63	Upper + lower extremities	17 (6-42)
Wang[30] 2001	72	39 (15-74)	73	Long bones of the upper + lower extremities	NR (NR)
Xu[46] 2009	69	38 (22-72)	64	Long bones of the upper + lower extremities	12,5 (6-84)
Beutler[36] 1999	27	35 (19-72)	NR	Long bones of the upper + lower extremities	9 (6-16)
Vogel[26] 1997b	48	38 (12-81)	52	Lower extremities	12 (6-48)
Schoellner[31] 2002	43	39 (18-74)	53	Long bones of the lower extremities	13 (9-51)
Rompe[28] 2001	43	40 (18-74)	53	Long bones of the lower extremities	11 (9-36)
Elster[35] 2010	192	45 (16-90)	73	Tibia	16,8 (NR)
Haffner[47] 2016	58	48 (16-82)	76	Tibia	15,6 (9-56)
Wang[50] 2009	42	35 (16-68)	52	Diaphysis of the long bones of the lower extremities	15,0 (6-48)
Kuo[42] 2015	22	30 (18-45)	59	Femoral shaft	10,5 (6-16)
Studies with undefined definition of delayed-union and nonunion					
Moretti[33] 2009a	204	NR (NR)	NR	Upper + lower extremities	NR (NR)
Valchanou[13] 1991	82	28 (9-76)	90	Upper + lower extremities	20,2 (NR)
Bara[32] 2007	81	NR (12-89)	68	Upper +lower extremities	8 (4-204)
Czarnowska-Cubala[40] 2013	31	47 (21-72)	65	Long bones of the upper + lower extremities	NR (NR)
Moretti[34] 2009b	10	NR (20-29)	100	Lower extremities	NR (NR)
Alvarez[37] 2011	34	50 (16-75)	22	Proximal metatarsal or zone 2/3 of the fifth metatarsal	6,8 (2,3-192,2)
West[45] 2008	28	48 (16-75)	21	Proximal metatarsal or zone 2/3 of the fifth metatarsal	13,3 (2,3-19,2)

DU: Delayed-union; NU: Nonunion; NR: Not reported; G: General anesthesia; R: Regional anesthesia; L: Local anesthesia; NA: No anesthesia; a: excluded from results due to <10 patients

Percentage previously infected nonunions	Type of nonunion in percentages			Type of anesthe-sia	Number of shocks applied	Energy density in mJ/mm²
	Hyper-thropic	Atro-phic	Oligo-trophic			
NR	10	17	73	L (n=17) NA (n=126)	3-5x 2500-3000	0,25-0,84
15	NR	NR	NR	R (n=51) NR (n=1)	3000	0,6
Excluded	NR	NR	NR	NR	1-3x 2000-4000	26 kV
NR	25	30	45	G or R	3500 – 10.000	0,3-0,4
NR	53	18	29	G or R	1000-6000	0,47-0,62
NR	84	16	NR	R or L	3000-10.000	0,56-0,62
NR	59	41	NR	NA(n=24) NR (n=1)	2x 2000	18 kV
17	NR	NR	NR	R (n=47) NR (n=1)	3000	0,6
NR	NR	NR	NR	R	3000	0,6
NR	NR	NR	NR	R	3000	0,6
21	38	41	NR	G or R	1-4x 2000-12.000	0,38-0,40
31,8	34,6	34,6		G or R	3000-4000	0,4
NR	83	17	NR	G	6000	0,62
NR	0	100	0	G or R	3000	0,58
NR	NR	NR	NR	NR	4000	0,22-1,10
NR	NR	NR	NR	R	1000 – 4000	1000 – 1700 bars
NR	NR	NR	NR	NA	1500-3000	500 bars
NR	NR	NR	NR	NR	3000	300 bars
NR	NR	NR	NR	NA	3x 4000	0,09-0,17
NR	NR	NR	NR	G or R	2000	0,22-0,51
NR	NR	NR	NR	G with regional block	2000	24 kV

Primary outcome: bony union

Studies at moderate risk of bias

The union rates reported by Cacchio et al.[8] after six months were 70% for ESWT group 1, 71% for ESWT group 2 and 74% for the surgical group. Union rates were not significantly different between the groups (X^2=0.08; p=0.95).

Studies at high risk of bias

The RCT of Zhai et al. [48] reported callus formation after six months in 55% of the patients who only received ESWT, and in 63% of the patients in the hBMSCs + ESWT group.

In the non-randomized trials, Notarnicola et al.[10] reported union rates at six months of 79% in the ESWT group and 78% in the surgical group. Union rates between the two groups were not significantly different (X^2=0.01; p=0.89). Furia et al.[9] reported union rates of 91% in the ESWT group and 90% in the surgical group after six months. No statistical analysis was done in this study.

The union rates that were reported in the 26 cohort studies, are shown in figure 2, and vary between 39%-100%. The overall union rates of all studies at high risk of bias are presented in table 5.

Table 5. Overall union rates of studies at high risk of bias

	Union rate (%)	**Total number of treated patients**
Delayed-unions treated with ESWT	86	314
Nonunions treated with ESWT	73	1782
Nonunions treated with surgery	81	80
Nonunions treated with hBMSC's and ESWT	62,5	32

Secondary outcome: adverse events

Cacchio et al.[8], Notarnicola et al.[10] and Furia et al. [9] compared adverse events between ESWT-treated patients and surgery-treated patients. The absolute number of complications are shown in table 6 and the overall complication rates in figure 3. The RCT of Zhai et al.[48] did not register adverse events.

Of the 26 included cohort studies, 23 studies registered adverse events after ESWT, [13, 25-37, 39, 41-47, 49] treating a total of 2027 delayed-unions and nonunions. Eight studies reported that no adverse events occurred after ESWT.[13, 32, 34, 36, 41, 42, 45, 47] Fifteen studies reported adverse events such as petechiae, local edema and hematoma's,[25-31, 33, 35, 37, 39, 43, 44, 46, 49] which are all grade 1 complications.

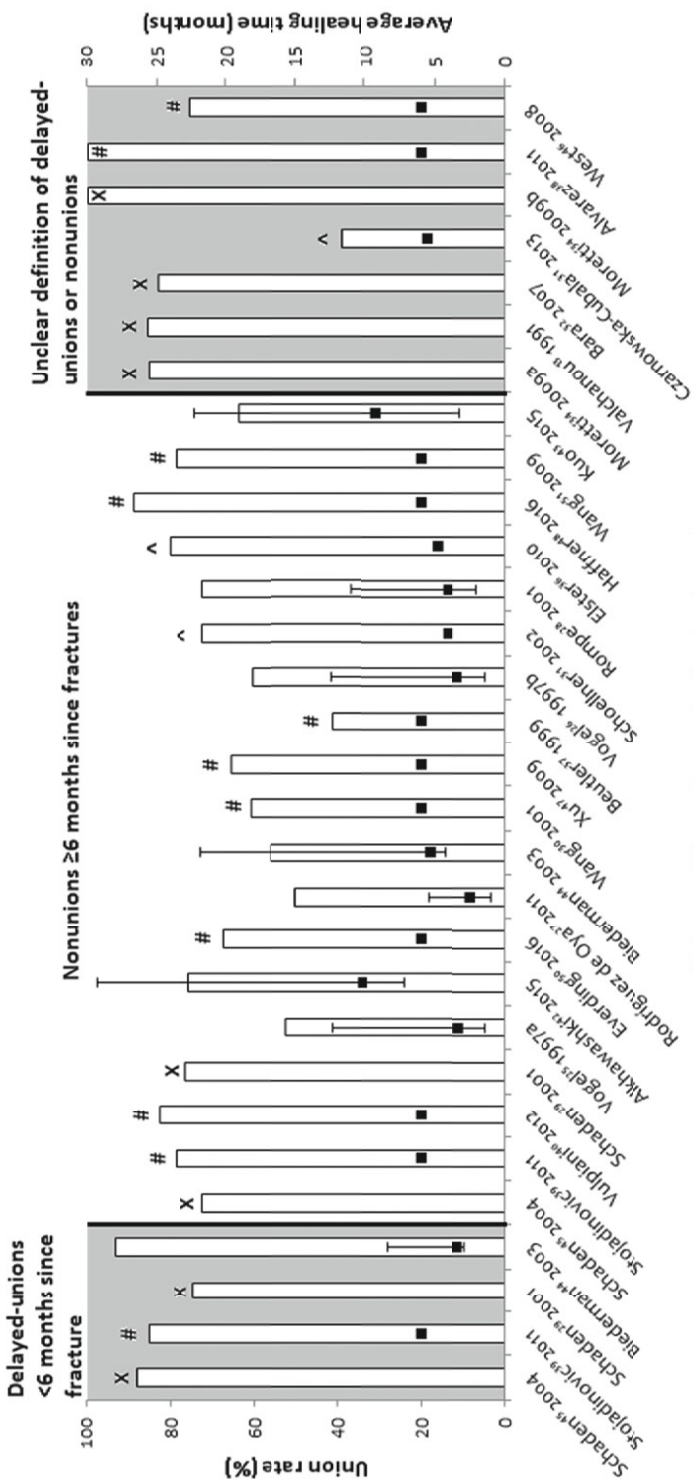


Figure 2. Union rates on the left x-axis and average healing times in months on the right x-axis reported in the 26 cohort studies. Error bars indicate the ranges in time to union after ESWT; # = study determined union rates at a fixed point of 6 months after ESWT; X = study did not report average union time; ^ = study did not report range in union time.

Table 6. Absolute number of complications classified by the adapted Clavien-Dindo classification

	ESWT				Surgery			
	Number of patients	Number of complications			Number of patients	Number of complications		
		Grade 1	Grade 2	Grade 3		Grade 1	Grade 2	Grade 3
Cacchio[8] 2009	84	23[1]	0	0	42	1[2]	0	2[3]
Notarnicol[10] 2010	58	0	0	0	60	0	0	0
Furia[9]2010	23	1[4]	0	0	20	1[5]	1[6]	9[7]
Overall	165	24	0	0	122	2	1	11

1. hematomas; 2. nerve neuropraxia; 3. wound infections requiring surgical debridement and antibiotics; 4. mild petechiae; 5. superficial cellulitis; 6. refracture requiring five weeks of immobilization in a walking boot; 7. hardware removal due to symptoms related to hardware

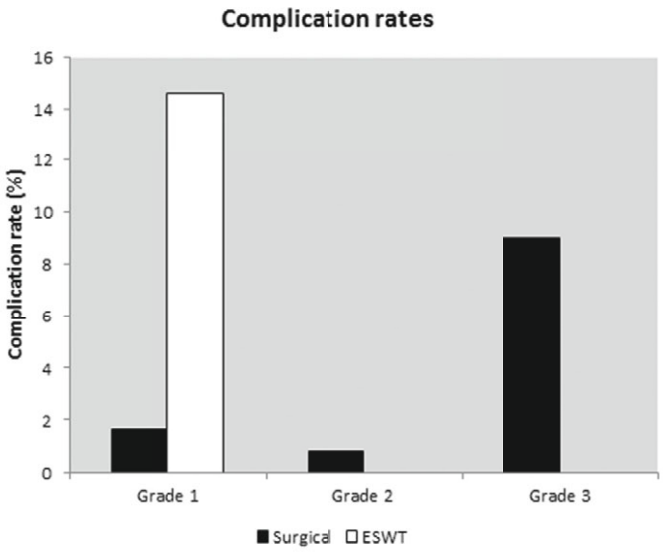


Figure 3. Overall complication rates based on the reported compilations in the studies of Cacchio et al.[8], Notarnicola et al.[10], and Furia et al.[9], classified by the adapted Clavien-Dindo classification[23].

Discussion

In this systematic review the effectiveness of ESWT in delayed-unions and nonunions was examined. The study of Cacchio et al.[8] showed that ESWT is as effective as surgical treatment for patients with long-bone nonunions, with unions rates between 71% and 74% after six months.[8] Next to this study, two more studies were published, in which

ESWT treatment was compared to a surgery treated control group. In concordance with the findings of Cacchio et al.,[8] both studies did find similar union rates between ESWT-treated patients and surgery-treated patients.[9, 10] The results of these studies seem to indicate that ESWT is as effective as surgery in the treatment of nonunions. The RCT of Zhai et al. [48] showed that hBMSCs transplantation with ESWT is more effective than ESWT alone, which shows that ESWT might be more effective with a combined treatment. These promising results are further supported by the included cohort studies, which together treated more than 2000 delayed-unions and nonunions and reported similar union rates as after surgery.

However, although we were able to identify 30 studies concerning this topic, the overall quality of those studies was poor, due to high risks of bias within the studies. The RCT of Cacchio et al. [8] was at moderate risk of bias and the results should therefore be interpreted with caution. The remaining 29 studies were all assessed as high risk of bias, due to missing control groups or non-randomized allocation to control groups, no blinding of the outcome assessors and participants, and unclear handling of incomplete data. Therefore, it is unadvisable to draw strong conclusions from these study results.

In addition to effectiveness, we also aimed to investigate the safety of ESWT treatment for nonunions. Twenty-three of the 30 studies addressed adverse events, treating together more than 1500 delayed-unions and nonunions. None of those studies reported any serious adverse events after ESWT, whereas severe adverse events were reported after surgery. None of the adverse events reported after ESWT needed further treatment. Based on these results it seems that ESWT is a safer treatment option for delayed-unions and nonunions than surgery.

Zelle et al.[20] published a systematic review on the effectiveness of ESWT in 2010 based on 10 studies. They suggested that approximately 75% of delayed-unions or nonunions could be treated successfully with ESWT, but that evidence is rather low because all ten studies were cohort studies[20]. Since the review of Zelle et al.,[20] multiple studies have been published on the effectiveness of ESWT.[9, 27, 35, 37-42] However, even after the conduction of those studies, the level of evidence remains low.

This review encountered some challenges and limitations. Firstly, as RCT's are the golden standard to prove the effectiveness of a treatment, we decided to perform the risk of bias assessment with a tool for RCT's. However, our search resulted in only two RCT's, and therefore all non-randomized and cohort studies were judged as high risk of bias. However, we believe that by using this tool, the lack of well performed RCT's is clearly pointed out. It is argued that nonunions are a biological end-point in which no further

bone healing will occur and that therefore a control group is not necessary to prove the effectiveness of ESWT. However, no clear consensus does exist when this biological endpoint is reached. Marsh et al.[51] showed that in patients with nonunions at six months, 50% experienced spontaneous fracture healing within the next four months. Biederman et al.[43] compared these results with the results of several studies that applied ESWT, and concluded that ESWT does not seem to accelerate bone union compared to natural fracture healing in nonunions. Thus, the cohort studies without control group might wrongly attribute the natural fracture healing process to ESWT, even after six months of nonunion.[43]

Secondly, there was a lot of heterogeneity within and between the studies. Most studies included in this review, included patients with fractures of different bones and with different types of delayed-unions and nonunions. Furthermore, studies used different energy settings for ESWT, differed in the number of shock waves applied to a fracture and the number of ESWT sessions that were done. Also, the type of anaesthesia differed between the studies (ranging from no anesthesia at all to general anesthesia) which might influence the effectiveness of ESWT[52]. Overall, due to the heterogeneity between the studies and the poor quality of the studies, combining the results in a meta-analysis would not have empowered our conclusion. More research should be done with homogeneous groups and shock wave parameters to be able to make recommendations about optimal shock wave parameters for particular fractured bones.

Lastly, some of the included studies were published by the same research groups.[25, 26, 30, 33, 34, 50] Looking at the results of those studies, it seems that some participants might have been included in the analysis of more than one publication. Therefore, this review might overestimate the actual amount of subjects that have been treated with ESWT.

In conclusion, the union rates that have been presented in this review after ESWT were comparable to union rates after surgery, and no serious adverse events have been reported after ESWT. Therefore, it seems that ESWT is as effective as surgery for the treatment of delayed-unions and nonunions, with less severe complications. However, the quality of the studies was poor and therefore the evidence for the effectiveness of ESWT for treatment of delayed-unions and nonunions is weak. We therefore hope that in the near future high quality RCT's will be conducted on the effect of ESWT in nonunions. These studies are essential to potentially implement ESWT into standard care.

References

1. Bhandari M, Guyatt GH, Swiontkowski MF, et al. A lack of consensus in the assessment of fracture healing among orthopaedic surgeons. J Orthop Trauma 2002;16:562-6.

2. Hak DJ. Editorial on "Epidemiology of fracture nonunion in 18 human bones". Ann Transl Med 2017;5:S19.

3. Zura R, Xiong Z, Einhorn T, et al. Epidemiology of Fracture Nonunion in 18 Human Bones. JAMA Surg 2016:e162775.

4. Schottel PC, O'Connor DP, Brinker MR. Time Trade-Off as a Measure of Health-Related Quality of Life: Long Bone Nonunions Have a Devastating Impact. J Bone Joint Surg Am 2015;97:1406-10.

5. Brinker MR, Hanus BD, Sen M, et al. The devastating effects of tibial nonunion on health-related quality of life. J Bone Joint Surg Am 2013;95:2170-6.

6. Schaden W, Mittermayr R, Haffner N, et al. Extracorporeal shockwave therapy (ESWT)--First choice treatment of fracture non-unions? Int J Surg 2015;24:179-83.

7. Koutalos A, Varitimidis S, Dailiana Z, et al. Operative management of humeral nonunions. Factors that influence the outcome. Acta Orthop Belg 2015;81:501-10.

8. Cacchio A, Giordano L, Colafarina O, et al. Extracorporeal shock-wave therapy compared with surgery for hypertrophic long-bone nonunions. J Bone Jt Surg Ser A 2009;91:2589-97.

9. Furia JP, Juliano PJ, Wade AM, et al. Shock wave therapy compared with intramedullary screw fixation for nonunion of proximal fifth metatarsal metaphyseal-diaphyseal fractures. J Bone Joint Surg Am 2010;92:846-54.

10. Notarnicola A, Moretti L, Tafuri S, et al. Extracorporeal shockwaves versus surgery in the treatment of pseudoarthrosis of the carpal scaphoid. Ultrasound Med Biol 2010;36:1306-13.

11. Peters RM, Claessen FM, Doornberg JN, et al. Union rate after operative treatment of humeral shaft nonunion--A systematic review. Injury 2015;46:2314-24.

12. El Haj M, Khoury A, Mosheiff R, et al. Orthogonal double plate fixation for long bone fracture nonunion. Acta Chir Orthop Traumatol Cech 2013;80:131-7.

13. Valchanou VD, Michailov P. High energy shock waves in the treatment of delayed and nonunion of fractures. INT ORTHOP 1991;15:181-4.

14. Wang CJ, Wang FS, Yang KD. Biological effects of extracorporeal shockwave in bone healing: a study in rabbits. Arch Orthop Trauma Surg 2008;128:879-84.

15. Chen YJ, Wurtz T, Wang CJ, et al. Recruitment of mesenchymal stem cells and expression of TGF-beta 1 and VEGF in the early stage of shock wave-promoted bone regeneration of segmental defect in rats. J Orthop Res 2004;22:526-34.

16. Wang FS, Yang KD, Kuo YR, et al. Temporal and spatial expression of bone morphogenetic proteins in extracorporeal shock wave-promoted healing of segmental defect. Bone 2003;32:387-96.

17. Wang FS, Yang KD, Chen RF, et al. Extracorporeal shock wave promotes growth and differentiation of bone-marrow stromal cells towards osteoprogenitors associated with induction of TGF-β1. Bone & Joint Journal 2002;84:457-61.

18. Kearney CJ, Lee JY, Padera RF, et al. Extracorporeal shock wave-induced proliferation of periosteal cells. Journal of Orthopaedic Research 2011;29:1536-43.

19. Ingber DE. Cellular mechanotransduction: putting all the pieces together again. The FASEB journal 2006;20:811-27.

20. Zelle BA, Gollwitzer H, Zlowodzki M, et al. Extracorporeal shock wave therapy: current evidence. J Orthop Trauma 2010;24 Suppl 1:S66-70.

21. Higgens JPR GSe. Cochrane Handbook for Systematic Review of Interventions Version 5.1.0. [updated March 2011]. Available from http://handbook.cochrane.org.; 2011.

22. Marsh JL, Slongo TF, Agel J, et al. Fracture and dislocation classification compendium - 2007: Orthopaedic Trauma Association classification, database and outcomes committee. J Orthop Trauma 2007;21:S1-133.

23. Dindo D, Demartines N, Clavien PA. Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. Ann Surg 2004;240:205-13.

24. Sink EL, Leunig M, Zaltz I, et al. Reliability of a Complication Classification System for Orthopaedic Surgery. Clinical Orthopaedics and Related Research® 2012;470:2220-6.

25. Vogel J, Rompe JD, Hopf C, et al. High-energy extracorporal shock wave therapy (ESWT) for treatment of pseudarthrosis. Z Orthop Ihre Grenzgeb 1997a;135:145-9.

26. Vogel J, Hopf C, Eysel P, et al. Application of extracorporeal shock-waves in the treatment of pseudarthrosis of the lower extremity. Preliminary results. ARCH ORTHOP TRAUMA SURG 1997b;116:480-3.

27. Rodríguez De Oya R, Sánchez Benítez De Soto J, García Munilla M. Treatment of nonunion and delayed consolidation with extracorporeal shockwaves. Rev Ortop Traumatol 2001;45:466-71.

28. Rompe JD, Rosendahl T, Schöllner C, et al. High-energy extracorporeal shock wave treatment of nonunions. Clin Orthop Relat Res 2001:102-11.

29. Schaden W, Fischer A, Sailler A. Extracorporeal shock wave therapy of nonunion or delayed osseous union. Clin Orthop Relat Res 2001:90-4.

30. Wang CJ, Chen HS, Chen CE, et al. Treatment of nonunions of long bone fractures with shock waves. Clin Orthop Relat Res 2001:95-101.

31. Schoellner C, Rompe JD, Decking J, et al. High-energy ESWT for pseudarthrosis. Orthopade 2002;31:658-62.

32. Bara T, Synder M. Nine-year experience with the use of shock waves for treatment of bone union disturbances. Ortop Traumatol Rehab 2007;9:254-8.

33. Moretti B, Notarnicola A, Moretti L, et al. Bone healing induced by ESWT. Clin Cases Miner Bone Metab 2009a;6:155-8.

34. Moretti B, Notarnicola A, Garofalo R, et al. Shock Waves in the Treatment of Stress Fractures. Ultrasound Med Biol 2009b;35:1042-9.

35. Elster EA, Stojadinovic A, Forsberg J, et al. Extracorporeal shock wave therapy for nonunion of the tibia. J Orthop Trauma 2010;24:133-41.

36. Beutler S, Regel G, Pape HC, et al. Extracorporeal shock wave therapy for delayed union of long bone fractures - Preliminary results of a prospective cohort study. Unfallchirurg 1999;102:839-47.

37. Alvarez RG, Cincere B, Channappa C, et al. Extracorporeal shock wave treatment of non-or delayed union of proximal metatarsal fractures. Foot Ankle Int 2011;32:746-54.

38. Stojadinovic A, Potter BK, Eberhardt J, et al. Development of a prognostic naïve Bayesian classifier for successful treatment of nonunions. J Bone Jt Surg Ser A 2011;93:187-94.

39. Vulpiani MC, Vetrano M, Conforti F, et al. Effects of extracorporeal shock wave therapy on fracture nonunions. Am J Orthop 2012;41:E122-7.

40. Czarnowska-Cubała M, Gwodziewicz K, Studniarek M, et al. Predictive role of scintigraphy (BS) in bone union induction using extracorporeal shock wave treatment (ESWT). J Orthop 2013;10:70-3.

41. Alkhawashki HMI. Shock wave therapy of fracture nonunion. Injury 2015.

42. Kuo SJ, Su IC, Wang CJ, et al. Extracorporeal shockwave therapy (ESWT) in the treatment of atrophic non-unions of femoral shaft fractures. Int J Surg 2015.

43. Biedermann R, Martin A, Handle G, et al. Extracorporeal shock waves in the treatment of nonunions. J Trauma Inj Infect Crit Care 2003;54:936-42.

44. Schaden W, Fischer A, Sailler A. Extracorporeal shock-wave therapy from viewpoint of traumatology. J Mineralstoffwechsel 2004;11:40-5.

45. West DL, Hawkins BJ, Langerman RJ. The use of extracorporeal shock waves in the treatment of delayed unions and nonunions. Curr Orthop Pract 2008;19:218-22.

46. Xu ZH, Jiang Q, Chen DY, et al. Extracorporeal shock wave treatment in nonunions of long bone fractures. Int Orthop 2009;33:789-93.

47. Haffner N, Antonic V, Smolen D, et al. Extracorporeal shockwave therapy (ESWT) ameliorates healing of tibial fracture non-union unresponsive to conventional therapy. Injury 2016;47:1506-13.

48. Zhai L, Ma XL, Jiang C, et al. Human autologous mesenchymal stem cells with extracorporeal shock wave therapy for nonunion of long bones. Indian j orthop 2016;50:543-50.

49. Everding J, Freistuhler M, Stolberg-Stolberg J, et al. [Extracorporal shock wave therapy for the treatment of pseudarthrosis : New experiences with an old technology] German. 2016.

50. Wang CJ, Yang KD, Ko JY, et al. The effects of shockwave on bone healing and systemic concentrations of nitric oxide (NO), TGF-β1, VEGF and BMP-2 in long bone non-unions. Nitric Oxide Biol Chem 2009;20:298-303.

51. Marsh D. Concepts of fracture union, delayed union, and nonunion. Clin Orthop Relat Res 1998:S22-30.

52. Rompe JD, Meurer A, Nafe B, et al. Repetitive low-energy shock wave application without local anesthesia is more efficient than repetitive low-energy shock wave application with local anesthesia in the treatment of chronic plantar fasciitis. J Orthop Res 2005;23:931-41.

Chapter 7

Extracorporeal shock wave therapy (ESWT) for talocrural arthrodesis; a double-blinded randomized controlled trial

A. Willems, O.P. van der Jagt, S.E. Meuffels Reig, S. Baart, M. Bloembergen, S.M.A. Bierma-Zeinstra, D.E. Meuffels
Submitted

Abstract

Background: A debilitating and common complication after talocrural arthrodesis is delayed union or non-union.

Purposes: The aim of this study was to investigate whether extracorporeal shock wave therapy (ESWT) can reduce the number of delayed unions after talocrural arthrodesis, compared to patients after sham-ESWT.

Patients and Methods: This multicenter double blinded randomized controlled trial included patients indicated for talocrural arthrodesis between July 2015 and June 2020. Patients were randomized between the intervention (per-operative ESWT) or control (sham ESWT) group. During ESWT 4.000 shocks were applied at a frequency of 4HZ and an energy-focused density of 0.5 mJ/mm^2. The primary outcome was the number of delayed unions assessed on computed tomography (CT) twelve weeks after talocrural arthrodesis. Less than 30% fusion on CT was considered a delayed union. Secondary outcomes were percentage fusion at six and twelve weeks, number of non-unions at 26 weeks, and patient reported outcome measures. The primary analysis was performed with a logistic mixed model.

Results: Forty-one talocrural arthrodesis were included in 37 patients. Nineteen arthrodesis were allocated to the intervention group and 22 to the control group. The number of delayed unions in the intervention group was 3/19 (16%) and in the control group 6/22 (27%). This was not significantly different (Risk difference=0.11; 95% CI 0.43 – 9.42; p=0.38). Also, no significant differences were found for the secondary outcomes ($p>0.05$).

Conclusion: Although there were less delayed unions in the intervention group, the difference between the groups was not significantly different. We could therefore not prove that ESWT is an effective treatment to decrease number of delayed unions after talocrural arthrodesis. However, the promising results should encourage more clinical studies.

Level of evidence: Level I

Introduction

Talocrural arthrodesis is the gold standard treatment for end-stage ankle osteoarthritis[8]. Although movement in the talocrural joint after the arthrodesis is not possible, overall function of the ankle increases due to pain reduction and decreased disability[10].

Successful fusion after talocrural arthrodesis is highly important for a good clinical outcome[13]. However, fusion problems are the most common complication after talocrural arthrodesis[15]. In literature, non-union rates after talocrural arthrodesis range between 8%-15%[3, 8, 10, 20]. Patients with non-union report a worse quality of life and functional outcome than patients with fusion[13]. In case of non-union, revision surgery is advised, which is accompanied by high costs due to additional imaging, pre- and postrevision visits and the additional surgery [9]. Effective treatments to reduce non-union rates after talocrural arthrodesis should therefore be further investigated.

A treatment that may improve bone fusion after talocrural arthrodesis is extracorporeal shock wave therapy (ESWT). ESWT is a non-invasive therapy during which acoustic pressure waves are generated [18]. These pressure waves are characterized by a high amplitude and short rising time, followed by low-magnitude negative wave[18]. It has been hypothesized that the biomechanical forces caused by the pressure waves trigger biochemical responses. This is called mechanotransduction[5]. Several pre-clinical studies reported biochemical responses after ESWT, causing cells to release osteogenic and angiogenic growth factors[4, 11, 12, 16, 24]. Most clinical studies with ESWT investigated the effect of ESWT in non-union fractures. A randomized controlled trial (RCT) showed that ESWT is as effective as surgery for treatment of non-union fractures[2]. Furthermore, a systematic review showed that ESWT can induce fracture healing in about 70% of non-union fractures[26]. In this review, 30 studies were included, treating more than 2000 non-union fractures.

The aim of this RCT is to investigate whether ESWT can decrease the number of delayed unions after talocrural arthrodesis compared to control patients after sham-ESWT. Furthermore, differences in non-union rate, fusion percentage and patient reported outcomes were studied.

Patients and methods

Study design

We conducted a prospective double-blinded multicenter randomized controlled trial. The protocol of this study has been approved by the local ethics committee (MEC-2014-176) and was prospectively registered at the 'International Clinical Trial Registry Platform' (https://trialsearch.who.int/; registration number: NTR5347).

Study participants

Patients were recruited between July 2015 and June 2020 at two hospitals in Rotterdam, The Netherlands (Erasmus MC University Medical Center and Maasstad Hospital). Patients that were indicated for talocrural arthrodesis by their orthopedic surgeon were screened for inclusion. These were patients who experienced symptomatic talocrural OA, sometimes accompanied with postural deviations of the ankle, but without postural deviations in the foot or complaints of the adjacent joints. Eligibility criteria for study participations were: 18 years or older, good understanding of the Dutch language and willing to participate. Exclusion criteria were previous ipsilateral talocrural arthrodesis or inability to walk. Eligible patients were informed about the study by standardized written and oral information. If patients decided to participate in the study, informed consent was signed and baseline measurements were carried out.

Randomization and blinding

Randomization was performed on the day of the talocrural arthrodesis. Participants were randomized to postoperative ESWT or sham-ESWT. Randomization was done by an independent researcher, with computer generated randomization lists based on block randomization, with variable sizes of randomization blocks (between 2-6), in a 1:1 ratio, and stratification for operating surgeon and surgical technique (open or arthroscopic). In patients with malalignment >15 that could not be redressed, talocrural arthrodesis was performed with the open technique, otherwise arthroscopy was used. An ESWT-certified researcher (AW) was informed about the randomization outcome. The patient and orthopedic surgeon were blinded for randomization outcome. Patients who participated for the second time with the contralateral talocrural joint were only randomized before the first talocrural arthrodesis and received the opposite randomization outcome for the second talocrural arthrodesis.

Intervention

In consultation with the anesthesiologist, patients had general anesthesia or spinal anesthesia during talocrural arthrodesis. After closure of the wound, when patients still had anesthesia, ESWT or sham-ESWT was performed. To protect the incisions, they were covered with a transparent medical dressing (3M Tegaderm). Extra caution was paid to avoid air bubbles underneath the dressing, as shock waves are poorly transmitted through air. An ultrasonic gel was applied to the ankle, to serve as coupling gel. To keep the orthopedic surgeon blinded, he/she left the operating room. In all cases ESWT was given by the same person.

If patients were randomized to the ESWT group, ESWT was performed with an electrohydraulic shock wave device (Orthogold 280, MTS Europe GmbH, Konstanz, Germany). In total, 4000 shocks were applied, 1000 shocks from the lateral side, 1000 shocks from the medial side, and 2000 shocks form the anterior side. X-ray positioning and a laser pointer were used for correct targeting of the talocrural joint. Shocks were generated at a frequency of 4Hz with an energy-focused density of 0.5 mJ/mm^2. ESWT took approximately 20 minutes.

If patients were randomized to the control group, sham-ESWT was performed. The researcher pretended to apply ESWT. The sound of ESWT was played on a wireless Bluetooth speaker. Sham treatment also took 20 minutes.

All patients stayed in the hospital overnight. After-treatment was the same for all patients. The first six weeks patients received a non-weight bearing lower leg cast. This period was followed by a lower leg walking cast for six weeks during which loading was allowed. After six and twelve weeks both radiographs and CT were performed. In the event that fusion had occurred after twelve weeks, the cast was removed. In case of delayed union casting was extended until fusion occurred. This decision was made by a blinded orthopedic surgeon based on x-ray imaging, computed tomography (CT), and physical examination. If lack of fusion was still suspected after 26 weeks, a CT was made and patients were advised for revision surgery.

Bone fusion assessment

Six and twelve weeks after talocrural arthrodesis CT was performed (SOMATON, Siemens healthcare AG, Erlangen, Germany). Multiplanar reconstructions in three planes (sagittal, coronal and axial), with a slice thickness between 1.5 and 2 mm were made. Consolidation of the talocrural joint was assessed based on the sagittal and coronal planes. On each slice, the width of the joint space was measured. The widths of all slices were then

summed. Also, the widths of the fused parts of the talocrural joint were measured and summed (figure 1). Osteosynthesis material crossing the joint space was excluded from the measurements. The sum of fused parts was divided by the sum of joint widths and multiplied by 100 to calculate the percentage of fusion[25]. The fusion scores of the sagittal and coronal planes were averaged to gain the final outcome of percentage fusion. Talocrural arthrodesis was judged to be fused if the percentage of fusion was higher than 30%[25]. CT were anonymized and the same observer assessed all scans.

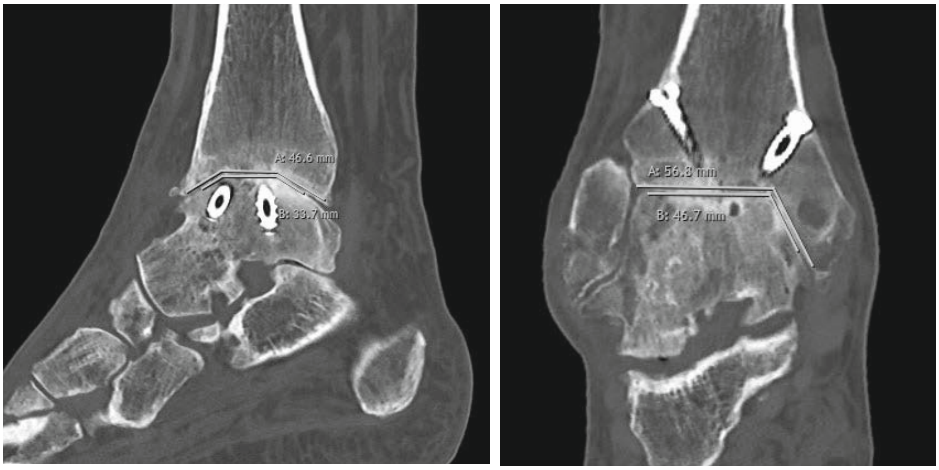


Figure 1. Sagittal and coronal CT images at twelve weeks with measurements of percentage fusion. Line A (upper line) representing joint space width and line B (lower line) representing fused part of talocrural joint.

A random set of five patients with CT at 6 and 12 weeks was re-assessed by a second blinded observer to assess reliability of the measurements. Interrater reliability scores were interpreted with the Koch-Landis method (kappa=0.01-0.2: slight agreement, kappa=0.21-0.40, fair agreement, k=0.41-0.60: moderate agreement, kappa=0.61-0.80: substantial agreement and k=0.81-1.00: excellent agreement)[14].

Outcome measures

The primary outcome measure was the number of delayed unions (≤30% fusion at CT after twelve weeks). Secondary outcomes were absolute percentages of fusion at six and twelve weeks, number of non-unions (≤30% fusion at CT after 26 weeks) and patient reported outcomes. Patient reported outcomes were the Foot and Ankle Outcome Score (FAOS), the American Orthopedic Foot and Ankle Score: Ankle-Hindfoot scale (AOFAS) and 36- item short form health survey (SF-36) [1, 6, 22]. FAOS and SF-36 were completed

by patients at baseline and at 16 and 26 weeks after talocrural arthrodesis. The AOFAS questionnaire has to be filled out partly by an orthopedic surgeon and was therefore only completed during standard clinical visits at baseline and 26 weeks. Outcome scores of FAOS, AOFAS and SF-36 range between zero and 100, where a higher score indicates better function or health.

Sample size

Sample size was calculated based on an intern analysis with patients who received an ankle arthrodesis in previous years in our center. This analysis showed that 40% of patients have a delayed union after talocrural arthrodesis. ESWT was expected to decrease delayed unions after talocrural arthrodesis to 10%. For sample size calculation, power was set to 80% and alpha at 0.05. To detect a difference in delayed unions, 30 patients were needed per group (total 60). To correct for potential loss to follow-up of 12% the target sample size was 68.

Statistical analysis

Descriptive statistics were checked for normality by visual inspection of the Q-Q plots and by performing a Shapiro-Wilk test. Variables with normal distribution were presented with mean and standard deviation (SD). In case of non-normal distribution variables were presented as median and interquartile range (IQR). Discrete variables were shown as counts and percentages.

Interrater reliability for consolidation measured on CT was analysed by a two-way random-effect model with absolute agreement. Interrater reliability was assessed on a random subset of 5 patients with a total of 10 CT's.

The primary outcome was analyzed with a logistic mixed model according to the intention to treat principle. In this model, delayed union was the dependent variable and treatment group the independent variable. The model was corrected for dependency in data of patients who participated with the left and right talocrural joints by adding a random effects term to the model. The same analysis was used for differences in number of non-unions at 26 weeks between intervention and control group.

Three level mixed models with random effects were used to assess differences in patient reported outcome measures. The model was corrected for dependency of data from patients with bilateral arthrodesis. A three level mixed model with random effects was

also used to assess differences in the amount of bone fusion at 6 and 12 weeks between intervention and control group.

Models were checked for linearity, homoscedasticity and normality of the residuals. Statistical significance was set at $p<0.05$. Data were analyzed by a blinded statistician using R Statistical software version 4.0.5.

Results

During the study period, 42 patients were eligible to participate in the study. Five patients refused to participate and therefore 37 patients were included and randomized. Four included patients were operated on both talocrural joints during the study period. Therefore, 41 talocrural joints of 37 patients were included in the study. Nineteen talocrural joints were treated with ESWT, and 22 with sham-ESWT. See figure 2 for a flowchart of the study participants. Due to technical problems with the shock wave device, one patient who was randomized to ESWT did not receive ESWT. For the primary outcome at twelve weeks, there was no lost to follow-up. See table 1 for the baseline characteristics of the study population.

Intrarater reliability

The interrater reliability for the percentage consolidation was excellent for the sagittal planes (k=0.96) and coronal planes (k=0.97). Also, the average consolidation between the sagittal and coronal planes was excellent (k=0.97).

Table 1. Baseline characteristics of the study population

	ESWT (n=19)	Control (n=22)
Age at inclusion	61 (11)	60 (9)
No (%) of men	63	73
Right side affected (%)	74	41
Smoker (%)	21	18
Body mass index	28 (4)	30 (5)
Surgical technique (%)		
Arthroscopic	63	77
Open	37	23

Table 1. continued part 2

	ESWT (n=19)	Control (n=22)
FAOS	34 (12)	40 (12)
AOFAS	35 (19)	46 (19)
SF-36		
- PhysFun.	33 (17)	36 (22)
- SocFun.	68 (24)	74 (30)
- LimPhys.	37 (36)	32 (37)
- LimEmo.	86 (28)	70 (43)
- Emo.	77 (13)	75 (20)
- Ener.	66 (17)	62 (19)
- Pain.	34 (17)	36 (19)
- GenHe.	69 (17)	63 (20)
- HeCha.	45 (20)	39 (25)

Data are mean (standard deviation) unless otherwise indicated. PhysFun.: Physical functioning; SocFun.: Social functioning; LimPhys.: Role of limitations due to physical health; LimEmo.: Role of limitations due to emotional problems; Emo.: Emotional well-being; Ener.: Energy/fatigue; GenHe.: General health; HeCha.: Health change

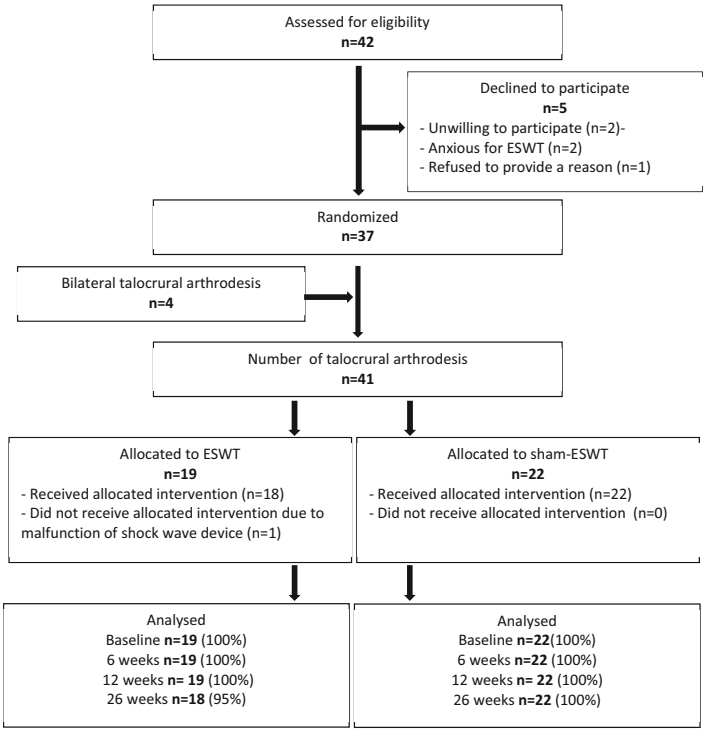


Figure 2. Flowchart of study participants.

Primary outcome

The number of delayed unions in the ESWT group were 3 out of 19 patients (16%) and in the control group 6 out of 22 patients (27%), a significant difference was not detected between the ESWT group and control group (Risk difference=0.11; 95% CI 0.43 – 9.42; p=0.38).

Secondary outcomes

The number of nonunions in the ESWT group were 3 out of 19 patients (16%) and in the control group 6 out of 22 patients (27%), a significant difference was not detected between the ESWT group and control group (Risk difference=0.11; p=0.43). Due to estimation difficulties with the data, CI could only be obtained via bootstrap method, resulting in exceedingly wide CI.

No significant differences were found for percentage fusion between ESWT and control group after six and twelve weeks, see table 2.

Patient reported outcome scores did not show any significant differences, see table 3.

Table 2. Results of percentage fusion

	ESWT	Control	Between group difference
6 weeks	45 (37 to 52)	39 (32 to 46)	-6 (-15 to 3)
12 weeks	52 (44 to 59)	46 (38 to 53)	-6 (-15 to 3)

Data are presented as mean estimate (95% confidence interval). 95% confidence intervals are based on logistic mixed models

Table 3. Results of patient reported outcome measures

	16 weeks			26 weeks		
	ESWT	**Control**	**Between group difference**	**ESWT**	**Control**	**Between group difference**
FAOS	47 (41 to 54)	47 (41 to 54)	0.09 (-9 to 9)	55 (49 to 62)	53 (46 to 60)	-2 (-11 to 6)
AOFAS				66 (58 to 75)	67 (59 to 75)	1 (-11 to 12)
SF-36						
- PhysFun.	37 (28 to 47)	38 (29 to 47)	0.5 (-12 to 13)	52 (42 to 62)	48 (38 to 57)	-5 (-17 to 8)
- SocFun.	52 (40 to 64)	66 (54 to 80)	14 (-0.4 to 28)	68 (55 to 80)	75 (63 to 86)	7 (-8 to 21)
- LimPhys.	19 (2 to 35)	23 (7 to 39)	5 (-18 to 27)	58 (41 to 75)	36 (20 to 53)	-21 (-44 to 2)
- LimEmo.	52 (35 to 71)	70 (53 to 87)	17 (-6 to 40)	66 (48 to 84)	82 (64 to 99)	16 (-8 to 39)
- Emo.	71 (64 to 77)	74 (68 to 81)	4 (-4 to 11)	77 (70 to 84)	77 (71 to 84)	0.4 (-7 to 8)
- Ener.	57 (50 to 65)	61 (54 to 69)	4 (-5 to 13)	64 (56 to 71)	65 (58 to 72)	1 (-8 to 10)
- Pain.	54 (43 to65)	52 (41 to 63)	-2 (-17 to 13)	60 (49 to 71)	60 (50 to 72)	1 (-14 to 16)
- GenHe.	65 (57 to 72)	69 (61 to 76)	4 (-4 to 12)	70 (62 to 77)	65 (57 to 72)	-5 (-13 to 3)
- HeCha.	56 (46 to 65)	44 (34 to 53)	-12 (-25 to 2)	60 (50 to 70)	48 (39 to 58)	-12 (-26 to 2)

Data are presented as mean estimate (95% confidence interval); PhysFun.: Physical functioning; SocFun.: Social functioning; LimPhys.: Role of limitations due to physical health; LimEmo.: Role of limitations due to emotional problems; Emo.: Emotional well-being; Ener.: Energy/fatigue; GenHe.: General health; HeCha.: Health change. 95% confidence intervals are based on linear mixed models.

Posthoc per protocol analysis

As one patient in the intervention group did not receive the intervention, we performed a per protocol analysis for the primary outcome with the one patient omitted from the analysis. The results of this analysis showed that number of delayed unions in the ESWT group were 2 out of 18 patients (11%) and in the control group 7 out of 23 patients (30%). A significant difference was not detected between the ESWT group and control group (Risk difference=0.19; 95% CI 0.63-19.50; p=0.15).

Bilateral arthrodesis

Four patients underwent bilateral arthrodesis, of which one side was treated with ESWT and one side with sham-ESWT. Two patients had no delayed unions on either side. Two patients had a delayed union on the sham-ESWT side, but not on the ESWT treated side.

Discussion

In this RCT we investigated whether ESWT can reduce the number of delayed unions after talocrural arthrodesis. The results of this study do not prove that ESWT is an effective treatment to significantly decrease number of delayed unions after talocrural arthrodesis. In the intervention group 3 out of 19 patients (16%) had delayed unions, whereas in the control group, 6 out of 22 (27%) patients had delayed unions. Although the number of delayed unions after talocrural arthrodesis was lower after ESWT, the difference was not significant and therefore effectivity of ESWT after talocrural arthrodesis could not be proven. However, looking at the absolute numbers from a clinical point of view, the decrease in number of delayed unions after ESWT is of interest and should be further investigated.

Our power calculation showed that 68 talocrural joints would be needed to show superiority of ESWT. Due to the worldwide COVID-pandemic, we were forced to stop the trial and where therefore not able to include the number of patients that were needed according to the power calculation. In our power calculation, we expected that 40% of patients would become a delayed union, and that this could be reduced to 10% with ESWT. However, the results of the current study showed that 27% of patients in the control group became a delayed union, and 16% in the ESWT group. This non-significant reduction of 11% is considerably less than the expected 30%. Therefore, even if the intended 68 talocrural joints would have been included, it is doubtful whether significant differences would have been found.

In this study, most patients were operated with arthroscopy. It has been shown that arthroscopically operated patients have higher fusion rates compared to patient operated with the open technique[17, 19]. Nielsen et al. (2008) reported fusion rates of 90% in arthroscopically treated patients versus 57% in patients operated with open technique[17]. The high numbers of arthroscopically treated patients may explain why the fusion rates in our study are higher than expected during the power calculation. The effect of ESWT after talocrural arthrodesis may be bigger in patients that are at increased risk for delayed-union, like patients treated with open surgery or smokers. However, our sample was too small to perform subgroup analysis.

This is the first RCT that investigated the effect of ESWT on bone fusion after talocrural arthrodesis. The study was conducted according to the CONsolidated Standards of Reporting Trials (CONSORT) 2010 guidelines[21], and therefore has a strong methodology. Risk of bias was low as a sham-treated control group was included in the study. Also, the orthopedic surgeons, patients, CT assessor and statistician were blinded for the

randomization outcome and there was no loss to follow-up. Although the differences between the ESWT group and control group were not significant, they are clinically relevant. The results strongly indicate a favorable effect of ESWT after talocrural arthrodesis. Looking at the results of this study it seems to be more likely that ESWT is an effective treatment than that it is not effective.

The study showed that the effect of ESWT on bone fusion after talocrural arthrodesis seems to be smaller than expected. However, even a small effect of ESWT after talocrural arthrodesis may be relevant as costs accompanied by non-union arthrodesis are high[9]. It would therefore be interesting for future research to study the cost-effectiveness of ESWT after talocrural arthrodesis. Based on the results of this study a new sample size calculation can be performed. With the observed results of this study, a power of 80% and alpha of 0.05, a total of 430 patients would be needed to reach significant differences in number of delayed unions. However, lower number may be needed to prove cost-effectiveness or in patients at increased risk of developing a delayed union.

Our secondary outcomes showed no significant differences. Non-union rates were the same as delayed union rates. At three months, a total of nine patients had a delayed union. In only one patient consolidation progressed to a fused talocrural joint. Remarkably, one patient with a fused talocrural joint at three months, became non-union at six months. These were both patients in the control group, and therefore absolute numbers of non-union did not change between three and six months. The percentages of fusion at six and twelve weeks were very similar between the intervention and control group, as were the scores of the patient reported outcome measures.

Based on the results of this study, it seems that ESWT is not as effective as expected in decreasing the number of delayed unions after talocrural arthrodesis. So far, a systematic review showed that ESWT seems to be effective for the treatment of non-union fractures. However, it must be noted that most studies included in this review were cohort studies without control groups[26]. Recently a retrospective cohort study with control group was published. This study showed that in patients with long bone non-unions, union rates are significantly higher if nail dynamization is combined with ESWT (88%) compared to nail dynamization alone (60%) [23]. To our knowledge, only one RCT investigated the effect of ESWT in non-union fractures. In this study with 126 patients, the union rate in the ESWT group was 70%, and in the surgical treatment group 73%. This study concluded that ESWT is as effective as revision surgery in the treatment of non-union fractures but with less serious complications[2].

In 2020 a retrospective cohort study was published on the effectiveness of ESWT for the treatment of arthrodesis non-unions[7]. In this study, fusion rates of arthrodesis in different body regions were analyzed. Union rates were highest in non-union arthrodesis in the hand (80%). Lower fusion rates were reported for arthrodesis of the talocrural joint (50%), subtalar joint (27%) and midfoot (0%). It seems that effectiveness of ESWT may differ between joints[7]. However, this study also had a small sample size (n=24)[7].

In conclusion, our data do not prove that ESWT is an effective treatment to decrease number of delayed unions after talocrural arthrodesis. However, given the results of this study it seems very likely that ESWT can decrease the number of delayed unions after talocrural arthrodesis. As the expected effect of ESWT seems to be smaller than initially thought, a larger study would be needed to investigate (cost-) effectiveness. This might be especially interesting to investigate in patients who are prone for developing a delayed union.

References

1. Aaronson NK, Muller M, Cohen PDA, Essink-Bot M-L, Fekkes M, Sanderman R, Sprangers MAG, te Velde A, Verrips E. Translation, Validation, and Norming of the Dutch Language Version of the SF-36 Health Survey in Community and Chronic Disease Populations. Journal of Clinical Epidemiology. 1998;51:1055-1068.

2. Cacchio A, Giordano L, Colafarina O, Rompe JD, Tavernese E, Ioppolo F, Flamini S, Spacca G, Santilli V. Extracorporeal shock-wave therapy compared with surgery for hypertrophic long-bone nonunions. J Bone Joint Surg Am. 2009;91:2589-2597.

3. Chalayon O, Wang B, Blankenhorn B, Jackson JB, 3rd, Beals T, Nickisch F, Saltzman CL. Factors Affecting the Outcomes of Uncomplicated Primary Open Ankle Arthrodesis. Foot Ankle Int. 2015;36:1170-1179.

4. Chen YJ, Wurtz T, Wang CJ, Kuo YR, Yang KD, Huang HC, Wang FS. Recruitment of mesenchymal stem cells and expression of TGF-beta 1 and VEGF in the early stage of shock wave-promoted bone regeneration of segmental defect in rats. J Orthop Res. 2004;22:526-534.

5. d'Agostino MC, Craig K, Tibalt E, Respizzi S. Shock wave as biological therapeutic tool: From mechanical stimulation to recovery and healing, through mechanotransduction. Int J Surg. 2015;24:147-153.

6. de Boer AS, Tjioe RJC, Van der Sijde F, Meuffels DE, den Hoed PT, Van der Vlies CH, Tuinebreijer WE, Verhofstad MHJ, Van Lieshout EMM, Group AS. The American Orthopaedic Foot and Ankle Society Ankle-Hindfoot Scale; translation and validation of the Dutch language version for ankle fractures. BMJ Open. 2017;7:e017040.

7. Everding J, Stolberg-Stolberg J, Pützler J, Roßlenbroich S, Ochman S, Raschke M. Extracorporal shock wave therapy for the treatment of arthrodesis non-unions. Arch Orthop Trauma Surg. 2020;140:1191-1200.

8. Fischer S, Klug A, Faul P, Hoffmann R, Manegold S, Gramlich Y. Superiority of upper ankle arthrodesis over total ankle replacement in the treatment of end-stage posttraumatic ankle arthrosis. Arch Orthop Trauma Surg. 2021.

9. Gagne OJ, Veljkovic AN, Glazebrook M, Penner M, Wing K, Younger ASE. Agonizing and Expensive: A Review of Institutional Costs of Ankle Fusion Nonunions. Orthopedics. 2020;43:e219-e224.

10. Glazebrook M, Burgesson BN, Younger AS, Daniels TR. Clinical outcome results of total ankle replacement and ankle arthrodesis: a pilot randomised controlled trial. Foot Ankle Surg. 2020.

11. Huang HM, Li XL, Tu SQ, Chen XF, Lu CC, Jiang LH. Effects of Roughly Focused Extracorporeal Shock Waves Therapy on the Expressions of Bone Morphogenetic Protein-2 and Osteoprotegerin in Osteoporotic Fracture in Rats. Chin Med J (Engl). 2016;129:2567-2575.

12. Kobayashi M, Chijimatsu R, Yoshikawa H, Yoshida K. Extracorporeal shock wave therapy accelerates endochondral ossification and fracture healing in a rat femur delayed-union model. Biochem Biophys Res Commun. 2020;530:632-637.

13. Krause F, Younger AS, Baumhauer JF, Daniels TR, Glazebrook M, Evangelista PT, Pinzur MS, Thevendran G, Donahue RM, DiGiovanni CW. Clinical Outcomes of Nonunions of Hindfoot and Ankle Fusions. J Bone Joint Surg Am. 2016;98:2006-2016.

14. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33:159-174.

15. Lawton CD, Prescott A, Butler BA, Awender JF, Selley RS, Dekker Ii RG, Balderama ES, Kadakia AR. Modern total ankle arthroplasty versus ankle arthrodesis: A systematic review and meta-analysis. Orthop Rev (Pavia). 2020;12:8279.

16. Li B, Wang R, Huang X, Ou Y, Jia Z, Lin S, Zhang Y, Xia H, Chen B. Extracorporeal Shock Wave Therapy Promotes Osteogenic Differentiation in a Rabbit Osteoporosis Model. Frontiers in Endocrinology. 2021;12.

17. Nielsen KK, Linde F, Jensen NC. The outcome of arthroscopic and open surgery ankle arthrodesis: A comparative retrospective study on 107 patients. Foot and Ankle Surgery. 2008;14:153-157.

18. Ogden JA, Tóth-Kischkat A, Schultheiss R. Principles of shock wave therapy. Clin Orthop Relat Res. 2001:8-17.

19. Quayle J, Shafafy R, Khan MA, Ghosh K, Sakellariou A, Gougoulias N. Arthroscopic versus open ankle arthrodesis. Foot Ankle Surg. 2018;24:137-142.

20. Ross BJ, Savage-Elliott I, Wu VJ, Rodriguez RF. Complications Following Total Ankle Arthroplasty Versus Ankle Arthrodesis for Primary Ankle Osteoarthritis. Foot Ankle Spec. 2021:1938640020987741.

21. Schulz KF, Altman DG, Moher D, the CG. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMC Medicine. 2010;8:18.

22. Sierevelt IN, Beimers L, van Bergen CJA, Haverkamp D, Terwee CB, Kerkhoffs G. Validation of the Dutch language version of the Foot and Ankle Outcome Score. Knee Surg Sports Traumatol Arthrosc. 2015;23:2413-2419.

23. Stolberg-Stolberg J, Fuchs T, Lodde MF, Roßlenbroich S, Garcia P, Raschke M, Everding J. Addition of shock wave therapy to nail dynamization increases the chance of long-bone non-union healing. J Orthop Traumatol. 2022;23:4.

24. Wang FS, Yang KD, Kuo YR, Wang CJ, Sheen-Chen SM, Huang HC, Chen YJ. Temporal and spatial expression of bone morphogenetic proteins in extracorporeal shock wave-promoted healing of segmental defect. Bone. 2003;32:387-396.

25. Willems A, Houkes CM, Bierma-Zeinstra SMA, Meuffels DE. How to assess consolidation after foot and ankle arthrodesis with computed tomography. A systematic review. Eur J Radiol. 2022;156:110511.

26. Willems A, van der Jagt OP, Meuffels DE. Extracorporeal Shock Wave Treatment for Delayed Union and Nonunion Fractures: A Systematic Review. J Orthop Trauma. 2019;33:97-103.

Chapter 8

General discussion

Aim of thesis

In this thesis three orthopedics topics were combined: osteoarthritis (OA), bone union assessment and extracorporeal shock wave therapy (EWST) to stimulate bone union. In the last chapter a randomized controlled trial (RCT) was presented in which patients with OA were treated with ESWT after talocrural arthrodesis to stimulate bone union. Therefore, it was important that all these topics were included in this thesis.

The first part of this thesis was about surgical treatment options for patients with knee and talocrural OA. Although surgical treatment of OA is daily practice, there are still knowledge gaps around these interventions. This thesis aimed to fill some of the knowledge gaps and focused on two unanswered questions concerning high tibia osteotomies (HTO) and talocrural arthrodesis. Thereafter, assessment of bone union was further explored. The aim was to gain more insight in the reliable and valid assessment of bone union. At last, the effect of ESWT on bone union was investigated. The aim was to assess effectiveness of ESWT in patients with talocrural arthritis to stimulate bone union after talocrural arthrodesis.

In this chapter we will discuss the generals results of our chapters, we will elaborate on limitations of this thesis and discuss future perspectives.

General results

Part 1: Osteoarthritis

We compared the accuracy of achieved correction between the medial open wedge HTO and the lateral closed wedge HTO. Based on our radiological results, both techniques can accurately correct varus malalignment, as the achieved corrections were not significantly different from the planned corrections. An important finding of our study was that there was no loss of correction in patients with the open HTO. A previous RCT did report loss of correction in open HTO[1]. However, in the previous study, open HTO were fixated with Puddu plates. These are plates with an incorporated wedge which are fixated with non-locking screws in four screw holes[2]. However, it has been shown that Puddu plates cannot rigidly fixate open HTO[2]. To overcome this fixation problem, angle-stable Tomofix plates were introduced in 2003[3]. Tomofix plates are designed according to the locking compression plate principle with locking screws to provide angular stability[2]. In our study, open HTO were fixated with angle-stable plates.

As our study showed no loss in correction in open HTO, angle-stable plates should be used in open HTO to achieve a stable fixation. Furthermore it was found that the closed HTO technique slightly overcorrected by 1.0° and the closed HTO technique slightly undercorrected by 0.9°. These over- and undercorrections were significantly different between the groups. For future studies it would be interesting to investigate whether over- or undercorrection makes a difference on (long-term) results, like progression of medial (or lateral) knee OA. Also, to truly know which technique is best, other radiological outcomes and patient reported outcomes should be considered on long-term.

We also investigated the effects of talocrural arthrodesis on OA development in adjacent joints. From literature we know that osteoarthritis is present in adjacent joints after talocrural arthrodesis[4]. However, it was unclear if OA was pre-existing to the talocrural arthrodesis or if it developed postoperatively. Our study was unique compared to all previously performed studies as we had pre-operative and postoperative computed tomography (CT). This is a strong feature as most studies lack pre-operative CT imaging[4]. Also, all studies with pre-operative imaging used radiographs to assess adjacent joint OA, which have been shown to lack bony detail for accurate OA assessment[5, 6]. Our study showed that OA was not increased before the talocrural arthrodesis. After talocrural arthrodesis, OA in adjacent joints was increased, compared to pre-operative CT and to the contralateral healthy control. Therefore, OA seems to develop as a consequence of talocrural arthrodesis. The development of OA after talocrural arthrodesis is probably caused by increased use and higher forces in the joints[7, 8]. However, it also seems that the increase in OA does not negatively affects patient reported outcomes and patient satisfaction. A limitation of this study was the relatively small sample size and the limited follow-up time of 7 years on average. To strengthen the conclusions of this study, it would be interesting to expand the number of patients within the cohort, by adding for instance patients from other clinics. With a higher sample size it would also be possible to analyze subgroups. For instance, development of arthritis in adjacent joints may be different for patients with rheumatoid arthritis or post-traumatic arthritis. However, we are not aware of a cohort of talocrural arthrodesis with pre-operative CT in other clinics. It would also be interesting to re-assess our cohort in the future to show the long-term effects of talocrural arthrodesis on OA in adjacent joints. This study showed us that patients develop adjacent joint OA after talocrural arthrodesis. Patients who are candidate for talocrural arthrodesis should be informed about development of asymptomatic OA in de adjacent joints. As clinical impact of OA development of adjacent joints seems to be limited, OA development should not be a reason to not perform talocrural arthrodesis.

Part 2: Bone union assessment

Bone union assessment is a commonly performed task by medical specialists like orthopedic surgeons or radiologists. However, bone union assessment is a rather subjective task as there is a lack of consensus[9]. The systematic review of Corrales et al. (2008) showed that in 123 studies investigating long bone-fractures, 11 different criteria were used to assess fracture healing[10]. As bone union was the primary outcome in our RCT presented in chapter 7, we were interested in an objective and valid method to assess bone union after talocrural arthrodesis. Furthermore, during the conduction of our systematic review on effectiveness of ESWT in delayed- and nonunions (chapter 6), one of the main challenges was the heterogeneity of methodology for bone union assessment. The different methodologies and criteria that were used for the assessment of bone union made it challenging to compare results between studies. Furthermore, we do not know if the percentage of bony union across de arthrodesis is related to clinical complaints and if there is a threshold when the relative amount of union reflects a stable, painless arthrodesis.

To decrease heterogeneity between studies and increase validity of bone union assessment, other research groups developed tools for bone union assessment. For example the Radiographic Union Scale in Tibial fractures (RUST) score for assessment of tibial fractures and REBRONE scale for assessment of long bone fractures[11, 12]. However, these tools did not seem appropriate to use in our RCT about effectiveness of ESWT after talocrural arthrodesis, as both instruments assess bone union with callus formation, thus secondary bone healing. After talocrural arthrodesis primary bone healing occurs, without the formation of callus. No tools exist that assess bone union after primary bone healing. Other limitations of the RUST and REBORNE tools are that they have no validated threshold to distinguish between fused and non-fused fractures. The REBORNE scale has a threshold for radiological consolidation but this threshold is not validated[12]. A threshold is clinically important as this can help clinicians and researchers to decide when a fracture, arthrodesis or osteotomy can be loaded. If loading starts to early, a fracture may displace or failure of osteosynthesis material may occur. However, unnecessary prolongation of the unloading period leads to stiffness, decreased muscle mass and productivity loss of the patient[13]. Lastly, the RUST score was developed to assess bone union from radiographs, and should therefore not be used to assess bone union from CT.

By performing two systematic reviews about bone union assessment, we strived to get closer to valid and reliable assessment of bone union. The first aim was to find a CT outcome parameter that is representative for bone union. The systematic review in pre-clinical studies showed that callus density and torsional rigidity were the most promising parameters to represent actual bone union. As we were primarily interested in assessing

bone union after talocrural arthrodesis, callus density would not be a representative parameter to assess bone union in our RCT. Torsional rigidity is a parameter that can be calculated from finite element analysis. In this analysis, a computer model of the bone is made based on CT-derived parameters such as bone mineral density, cross section of the bone, and shape of the bone. From this model, the torsional rigidity can be calculated. Associations between actual bone union and other parameters, like bone mineral density, bone volume and callus volume, were less strong, and conflicting between the studies.

Thus, for our purposes, torsional rigidity seemed to be the most appropriate parameter for bone union assessment after arthrodesis. However, this parameter has only been investigated in pre-clinical studies with secondary bone healing and should therefore first be further investigated in models with primary bone healing. Also, advanced software and knowledge is needed to conduct analysis to assess torsional rigidity and bone mineral densities must be assessed accurately for the analysis with phantoms. As our CT protocol did not include phantoms, this information was not available within our study.

In addition to the pre-clinical systematic review on bone union, we also conducted a systematic review of clinical studies. Within this review we focused on assessment of bone union from CT after foot and ankle arthrodesis, because this was the main reason to investigate this topic. Interestingly, none of the clinical studies that used CT for bone union assessment used any of the CT parameters described in our pre-clinical review. This is probably because the parameters mostly assess secondary bone healing, are not able to accurately assess bone union (like bone mineral density, bone volume and callus volume) or are not widely available for clinical use (like torsional rigidity). However, it does also shows that there is quite a big gap between pre-clinical and clinical research. In our clinical review, we showed that calculating percentage fusion is probably the most valid and accurate currently available method for bone union assessment. This is done by summing the fused parts and the widths of the talocrural joint on the sagittal and coronal slices. The sum of the fused parts is divided by the sum of the widths and multiplied by 100. However, for future research it would be of added value if this methodology would be validated in a pre-clinical study.

The systematic reviews on bone union show that there are several interesting methodologies for bone union assessment. The most promising methods at this moment are calculating torsional rigidity using finite element analysis and calculating bone fusion. However, more research should be done to further validate these methods and to set cut-off values. Also, to increase scientific and clinical applicability of both methods, it would of high value if these parameters could be used in combination with artificial intelligence, through which bone union could be (semi)automatically assessed.

Part 3: Extracorporeal shock wave therapy (ESWT)

Another aim of this thesis was to investigate the effectiveness of ESWT to stimulate bone union. The systematic review that we conducted about the effectiveness of ESWT in delayed- and nonunion fractures showed that ESWT seems to be an effective treatment, but the quality of the included studies was low. Therefore, the evidence is weak and more good quality studies should be conducted. However, to date it seems very difficult to set-up collaborations with the industry or find financial support with (inter)national grants to perform such high quality clinical studies.

The RCT that we performed in chapter 7 about ESWT after talocrural arthrodesis was a methodologically strong study, which should result in a high level of evidence due to a low risk of bias. A strong feature of this study was that patients were randomized between ESWT and sham-ESWT. Patients, orthopedic surgeons, CT assessor and statistician were blinded for the randomization outcome. However, we were not able to reach the calculated power and therefore it remains difficult to draw strong conclusions from this study. Although we were not able to find significant differences between the groups, looking at the absolute values it seemed that there was a trend for less delayed unions in the ESWT group. With more power we might have found significant differences between the groups. Based on this study we cannot prove that ESWT is an effective therapy to stimulate bone union after talocrural arthrodesis.

Limitations

Feasibility of studies

In this thesis, two RCT's are presented, which should be providing high level evidence (chapters 2 and 7). However, both studies did not reach the intended power as the anticipated number of patients were not included. In both studies, the reason for inclusion problems were similar. Eligible patients were generally willing to participate in the studies, but both RCT 's struggled with low numbers of eligible patients. Both studies were eventually terminated during the worldwide COVID-pandemic as the pandemic caused a further decrease in inclusion rate and elective surgeries (including talocrural arthrodesis and HTO) were postponed.

Literature shows that more studies struggle with lack of power. Abdullah et al. (2015) investigated orthopaedic studies that did not find a significant difference in their primary outcome. They showed that almost 30% of these Orthopedic studies were underpowered[14]. More than half of these studies did not perform an a priori power

calculation[14]. Lack of power may lead to type II errors, meaning that a study does not find a significant difference although a difference does exist. Type II errors may lead to wrong conclusions and clinical decisions.

Quickly after start of our ESWT RCT (chapter 7) we noticed that inclusion rate was problematic. Therefore we tried to increase inclusion numbers by generating more awareness of the study. We presented the study at a meeting of the Dutch Foot and Ankle Society, to reach out to Orthopedic foot and ankle surgeons within the Netherlands. The aim of the presentation was to increase inclusions by either convincing orthopedic surgeons to refer their patients to our clinic or by expanding the study to other clinics. If every clinic within our region would have referred one or two patients to our clinic we probably would have been able to reach the intended power. However, orthopedic surgeons were not eager to refer their patients to our clinic, probably due to loss of income and commitment to their own patients. Some clinics showed interest to participate in the RCT as a study center. However, only one clinic proved to have high numbers of eligible patients. Therefore, the study was expanded to this clinic. Initially, this led to an increase of the inclusion rate. However, changes in the staff of the center led to a severe decrease of inclusion rate.

In our study, the complete process of finding an eligible center and gaining local medical approval took more than one year, in which valuable study time was lost. For future RCT's we would strongly advise to carefully evaluate the number of eligible patients before the start of the study. In case of our RCT's this could have been done by checking surgical codes over the last years in the electronic patient record. Collaborations between hospitals should be stimulated to reach the intended power and studies should be conducted in high-volume centers. These recommendations may seem obvious. However, in practice it can be challenging to do so. For example, you may aim for a big multicenter RCT but the grant does not provide enough financial support this, or other centers are not willing to participate because participating is too time consuming. Some of these problems may be solved by a study design called small simple trials (SST), which was proposed by wright et al. (2018)[15]. An SST is basically a simplified RCT, in which only essential data are collected with as few visits as possible. The study question of an SST should be uncomplicated, eligibility criteria should be broad, and data collection limited. The outcome measure should be objective and should be assessed on short-term. Also, the outcome should preferably be measured on a continuous level as continuous variables provide higher power than binary variables[15]. To gain high-level evidence, the strong aspects of an RCT must be maintained, like clear eligibility criteria, randomization, blinding and sample size calculation.

When applying a SST design to our RCT on ESWT presented in chapter 7, our study design would have been different. Our primary aim was to assess differences in the number of delayed unions after talocrural arthrodesis in patients with treated with ESWT or sham-ESWT. Instead of using this binary outcome, we could have chosen for a continuous outcome for example the percentage of fusion. Also, during the study, three CT's were made, to monitor fusion progression. However, for our primary aim we only needed one CT after three months. The costs that could have been saved by conducting less CT's could have been used to add more participating centers. The savings could have been used for transportation and insurance costs of the shock wave device or to buy/lend another shock wave device. Patient reported outcomes can be collected relatively easy with digital questionnaires and are not a (financial) burden. However, one patient reported outcome that was used (the American Orthopedic Foot & Ankle Society (AOFAS) Ankle-Hindfoot Scale) must be partly filled out by a physician based on physical examination. This is an extra study task for participating physicians, which is not absolutely necessary. Also, the added value of the AOFAS scale is very limited as we already asked patients to complete the Foot and Ankle Outcome Score (FAOS). Therefore, the AOFAS could have been removed from the study design to decrease the study burden for the physicians and patients.

With this simplified study design, it would have been more likely that the intended power would have been reached in our ESWT RCT. Firstly, because of the continuous outcome. Secondly because it would have been more easy to include other centers as the study burden is very limited and transportation of the shock wave device could have been afforded within the study budget. A disadvantage of this design is that information will be lost compared to the more extended design. For example, we would not have been able to monitor fusion progression due to a missing CT at six weeks. Also, the SST design may have found a difference in percentage fusion between the two treatment groups, but the clinical relevance of this difference may remain unclear.

However, the SST design is not applicable to all RCT's with inclusion problems. For the RCT about HTO's a SST design would probably not have led to more inclusions. In this study, patients were willing to participate. The HTO study received no funding and the study design was therefore quite basic, i.e. the study design did not include a lot of extra examinations or burden for physicians and patients. The recruitment problems were the result of low numbers of performed HTO's during the study. During the study set-up it was expected that more HTO's would be performed, but once the study started the HTO's seem to have disappeared. This is a well-known phenomenon in clinical research, which is called 'Lasagna's law'. Lasagna's law states that investigators and physicians usually overestimate the number of patients available for a study. It seems that as soon as a

study starts, the number of eligible patients becomes a fraction of what it was assumed to be[16]. It seems that it is very hard for investigator and physicians to estimate the number of eligible patients. This once more shows how important it is to check number of available patients before the start of a study in an electronic patient record.

Overall, RCT's usually have multiple aims, which make the study designs relatively complex. We sometimes seem to forget that answering the primary aim is the most important matter when performing a RCT. Therefore, when designing a study with possible power issues, we should keep the study as simple as possible to reach the intended power, avoid type II errors, and come to strong conclusions.

Valid outcome measures

A valid and reliable outcome measure is important for every study. Risk of bias increases enormously if validity or reliability of an outcome measure is insufficient. Studies with high risk of bias may lead to incorrect conclusions. It is therefore important to carefully consider which outcome measure should be used.

One of the main topics of this thesis was to find a valid and reliable method for bone union assessment. Despite our research, we were not able to find the perfect method. In our RCT on ESWT (chapter 7) bone union was calculated by the percentage fusion. Within this thesis we did assess the reliability of this method, by assessing inter-observer reliability. However, the validity of this method has not been extensively researched. To assess the validity of this method it should be compared to the golden standard for bone union assessment. It is debatable what the golden standard for bone union assessment is but we could look at associations with biomechanical or histological testing. For future studies, we strongly recommend to assess the validity of calculating percentage fusion.

In this thesis, the American Orthopedic Foot & Ankle Society (AOFAS) Ankle-Hindfoot Scale was used in two studies (chapters 3 and 7) [17]. This seems to be a logical choice as the AOFAS Ankle-Hindfoot scale is a commonly used tool. In studies about total ankle arthroplasty, it is even the most commonly used outcome measure[18]. In 2017, the AOFAS Ankle-Hindfoot Scale was translated and validated into the Dutch language [19]. However, over the years concerns have been raised about the reliability and validity of this outcome measure[20]. One of the problems with the AOFAS scale is that it must be partly filled out by a physician, based on physical examination. However, inter-observer scores were not assessed for these questions. It is also unclear whether the remaining questions should be filled out by the physician or by the patient. Furthermore, several studies investigated the validity of the score by investigating associations between the

AOFAS Scale and other questionnaires, like Short Form-36 (SF-36) and Musculoskeletal Function Assessment (MFA) [21, 22]. These studies concluded that there are low levels of correlation between AOFAS scale and SF-36 and MFA. The problems with reliability and validity of the AOFAS scale led to a statement of the AOFAS in 2011, in which they recommended that the OAFAS scale should not be used as an outcome measure due to insufficient reliability and validity[20]. However, despite this statement, the AOFAS scale is still widely used, like it was also used in our studies. We would therefore like to emphasize the importance of a reliable and validated outcome measure and strongly recommend to check the reliability and validity of outcome measures before using them for clinical research. In retrospect, the use of AOFAS Ankle-Hindfoot scale should have been avoided in chapters 3 and 7. The Foot and Ankle Outcome Score (FAOS), which was also included in these chapters, is a better subjective outcome score to represent foot and ankle complaints[23].

Future perspectives

Bone union assessment

In this thesis, we extensively elaborated on radiological parameters for bone union assessment. We encourage future studies to further investigate methods for valid and objective bone union assessment. For instance, before torsional rigidity and percentage fusion can be clinically used to assess fracture healing, they should be validated and a cut-off values must be determined to distinguish between fused and non-fused bones.

Ultimately, it should be possible to develop a (semi)automatically methodology to assess bone fusion through an artificial intelligence model. Artificial intelligence (AI) is a process in which a computer is used to model intelligent behavior needing minimal human intervention. Currently, no artificial intelligence models exist to assess bone fusion. Future machine learning models may use the above-described parameters to develop accurate models. Over the past years, applications with AI have majorly developed. AI has for instance been used for detection of bone cancer, OA and fractures[24-30].

Within the last ten years, several studies were published in which AI is used for fracture detection through deep learning. Deep learning is a form of artificial intelligence in which a computer model is able to make classifications based on images[31]. The novelty of deep learning is that the model itself can determine the best imaging features to differentiate between fractured and non-fractured bones[32]. This is a big step forward compared to earlier image-based machine learning which required human input to

determine which features (like for instance image intensity values or region of interest) are most important[32]. The deep learning models are trained by using large datasets and can achieve accuracy levels higher than accuracy of humans. Studies about deep learning and fracture detection have shown very promising results. Olczak et al. (2017) investigated the feasibility of deep learning for fracture detection from radiographs of the wrist, hand and ankle[26]. In this study, 256.458 skeletal radiographs were included of which 56% of the images showed fractures. The accuracy of fracture detection from artificial intelligence was 83%, which was similar to the accuracy of fracture detection of two senior orthopedic consultants[26]. Chung et al. (2018) evaluated the accuracy of deep learning for the detection of proximal humerus fractures and found an accuracy of 96%. This accuracy was statistically significantly higher compared to general physicians ($p<0.001$) who had an accuracy of 85%, and similar to general orthopedist and shoulder specialized orthopedics that both had accuracy scores of 93%[27]. The study of Lindsey et al. (2018) showed that the accuracy of diagnosing writs fractures is similar between deep learning and specialized orthopedic surgeons, and better compared to emergency clinicians[28]. Therefore, if emergency clinicians could use deep learning for fracture assessment, this would probably improve the accuracy of their diagnosis.

Although the results of these studies are very promising, some issues must be overcome before AI can be clinically used for fracture detection. The previously described studies used only single images for fracture identification, to keep the models relatively simple. Models based on series of images from different projections, or on CT, would improve the diagnostic performance of the model, but would also demand more memory of graphic processing units or increased model training time[26, 27, 29].

To our knowledge, no studies are available on AI and bone union assessment from radiographs or CT. We encourage the development of such an AI model to assess bone union, as a well working model will improve patient care and clinical research. This thesis does presents a thorough overview of the current literature on bone union assessment with CT. Parameters like torsional rigidity, callus density or percentage fusion may be used in future AI models for assessment of bone union.

ESWT

The initial reason for the thesis was to assess the effectivity of ESWT in bone union problems. This thesis showed that studies investigating bone union and ESWT are generally of low quality, but that many studies do show positive effects of ESWT on bone union. These are pre-clinical studies, as well as clinical studies[33-36]. Although our RCT (chapter 7) did not prove effectiveness of ESWT, the results seem to indicate that ESWT

may be effective to stimulate bone union. Also, there are no studies of high quality that shows that ESWT is not effective. If ESWT is effective for bone union problems, it could be applied for many different indications within orthopedics, like nonunion fractures of the long bones, union problems after arthrodesis, nonunions of scaphoid fractures, osteoporosis or atypical femur fracture (AFF) [37, 38]. An advantage of ESWT above surgical intervention in treatment of nonunion fractures is that ESWT is noninvasive, and that patients therefore are at low risk of complications[33]. Also, treatment of nonunions is expensive and may be less expensive with ESWT[39].

Within the scope of this thesis a pilot-study about the effect of ESWT on AFF was planned to be performed. For this study, all preliminary work was done, like approval of the Dutch Health and Youth Care Inspectorate for off-label use of the shock wave device and approval of the Medical Ethical Committee. However, just before the start of the study the ESWT device was sold to another hospital, making it impossible to perform this study. The aim of this pilot study was to investigate whether ESWT is an effective treatment for non-healing AFF. AFF are non-traumatic fractures of the femur that occur as a results of long-term bisphosphonate use in patients with osteoporosis. Bisphosphonates decrease bone resorption, therewith increasing the BMD and resulting in decreased fracture risk. However, it has been hypothesized that decreased bone resorption on long-term could decrease repair of micro fractures and result in accumulation of micro damage. This might eventually lead to a complete non-traumatic fracture of the femur[40]. Another problem within AFF is that fracture healing is complicated due to the long-term bisphosphonates usage, as this negatively affects bone union[41]. AFF have considerably lower union rates after surgery than regular femur fractures (54 vs 99%) [42] and in those patient with union, time to union is increased (10 months in AFF vs 4 months in regular femur fractures)[43-45]. Due to the high rate of nonunion and long union time, the osteosynthesis material is stressed for a very long period and failure of the material is common, resulting in refractures, inability to stand and walk and (multiple) revision surgery[45]. Due to the bone union problems in AFF, ESWT may be a very interesting treatment for these patients, as ESWT might be the extra impulse needed for bone union. Although this pilot-study was not performed within this thesis, it could be of high value for AFF patients if the study would be performed in the future.

We can give several reasons why we encourage more clinical trials about the effectiveness of ESWT on bone union. The first reason is the promising results of ESWT on bone union that were shown in our review (chapter 6) and RCT (chapter 7). Secondly, ESWT can be used for multiple indications and can therefore improve treatments of several bone (healing) disorders. It therefor has the potential to improve treatments of multiple orthopedic problems. Also, ESWT has been shown to be very safe. In this thesis no severe

complications were reported after ESWT. Therewith, we can state that ESWT is safer than surgical treatment. Lastly, due to the non-invasiveness of ESWT it is likely that ESWT is less expensive than surgical treatment.

For a future study on the (cost-)effectiveness of ESWT we would suggest to conduct this study on an indication that occurs frequently. This is dependent on the center performing the study, but could be nonunion fractures of long bones, scaphoid nonunions or arthrodeses. Also, in a future study a multicenter set-up from the start should be considered and a SST design should be considered. Also, we would recommend investigating cost-effectiveness of ESWT.

Conclusion

In this thesis, we addressed two knowledge gaps concerning OA related surgeries. We showed that in patients with medial knee OA and varus malalignment, malalignment can be accurately corrected by the open HTO technique as well as with the closed HTO technique. Also, in both techniques the fixation was stable, without loss of correction during two years of follow-up. Therefore, from a radiological point of view, both techniques can be used to successfully perform a HTO.

We also investigated the development of OA in patients who had had a talocrural arthrodesis. Our study showed that OA was not increased before talocrural arthrodesis, but that OA scores increased after talocrural arthrodesis. It therefore seems that talocrural arthrodesis causes OA development in adjacent joints. However, the degree of OA in adjacent joints was not related to patient reported outcomes and therefore the clinical impact of OA increase seems to be limited. Therefore, OA development should not be a reason to not perform a talocrural arthrodesis.

This thesis delved into the valid and accurate assessment of bone union. It provides a solid overview of the current literature and shows that torsional rigidity and percentage fusion are parameters, which could potentially be used for valid and accurate bone union assessment from CT. In the future, these parameters should be used in AI models for bone union assessment. For scientific as for clinical purposes it would be of high value if a valid, reliable and fast AI model could assess bone union. Hopefully, this thesis contributes to this goal by conducting some important preliminary work.

Although we could not prove effectiveness of ESWT for bone union indications, the results presented in this thesis do seem to indicate a favorable effect of ESWT on bone union. These promising results should encourage the conduction of more studies concerning this topic. Hopefully, ESWT may become a safe and effective treatment for many bone union indications in the future.

References

[1] R. W. Brouwer, S. M. Bierma-Zeinstra, T. M. van Raaij, and J. A. Verhaar, "Osteotomy for medial compartment arthritis of the knee using a closing wedge or an opening wedge controlled by a Puddu plate. A one-year randomised, controlled study," J Bone Joint Surg Br, vol. 88, no. 11, pp. 1454-9, Nov 2006.

[2] A. M. Ahmed, A. Addosooki, A. Saleh Sleem, and E. Said, "Superior survivorship and plate-related results of TomoFix compared to Puddu plate fixation for opening-wedge high tibial osteotomy: A systematic review of the literature," Knee, vol. 42, pp. 1-18, Feb 28 2023.

[3] A. E. Staubli, C. De Simoni, R. Babst, and P. Lobenhoffer, "TomoFix: a new LCP-concept for open wedge osteotomy of the medial proximal tibia--early results in 92 cases," Injury, vol. 34, pp. 55-62, 2003.

[4] J. S. Ling et al., "Investigating the relationship between ankle arthrodesis and adjacent-joint arthritis in the hindfoot: a systematic review," J Bone Joint Surg Am, vol. 97, no. 6, pp. 513-20, Mar 18 2015.

[5] D. Vier, T. Louis, D. Fuchs, C. T. Royer, J. R. Zide, and D. E. Jaffe, "Radiographic assessment of the subtalar joint: An evaluation of the Kellgren-Lawrence scale and proposal of a novel scale," Clin Imaging, vol. 60, no. 1, pp. 62-66, Mar 2020.

[6] D. J. Mayich, E. Pinsker, M. S. Mayich, W. Mak, and T. R. Daniels, "An analysis of the use of the Kellgren and Lawrence grading system to evaluate peritalar arthritis following total ankle arthroplasty," Foot Ankle Int, vol. 34, no. 11, pp. 1508-15, Nov 2013.

[7] R. Schuh, J. G. Hofstaetter, S. G. Hofstaetter, S. B. Adams, Jr., K. H. Kristen, and H. J. Trnka, "Plantar pressure distribution after tibiotalar arthrodesis," Clin Biomech (Bristol, Avon), vol. 26, no. 6, pp. 620-5, Jul 2011.

[8] Y. Wang, Z. Li, D. W. Wong, and M. Zhang, "Effects of Ankle Arthrodesis on Biomechanical Performance of the Entire Foot," PLoS One, vol. 10, no. 7, p. e0134340, 2015.

[9] J. S. Fisher, J. J. Kazam, D. Fufa, and R. J. Bartolotta, "Radiologic evaluation of fracture healing," Skeletal Radiol, vol. 48, no. 3, pp. 349-361, Mar 2019.

[10] L. A. Corrales, S. Morshed, M. Bhandari, and T. Miclau, 3rd, "Variability in the assessment of fracture-healing in orthopaedic trauma studies," J Bone Joint Surg Am, vol. 90, no. 9, pp. 1862-8, Sep 2008.

[11] D. B. Whelan et al., "Development of the Radiographic Union Score for Tibial Fractures for the Assessment of Tibial Fracture Healing After Intramedullary Fixation," Journal of Trauma and Acute Care Surgery, vol. 68, no. 3, 2010. [Online]. Available: https://journals.lww.com/jtrauma/Fulltext/2010/03000/Development_of_the_Radiographic_Union_Score_for.24.aspx.

[12] E. Gómez-Barrena, N. G. Padilla-Eguiluz, E. García-Rey, P. Hernández-Esteban, J. Cordero-Ampuero, and J. C. Rubio-Suárez, "Validation of a long bone fracture non-union healing score after treatment with mesenchymal stromal cells combined to biomaterials," Injury, vol. 51, pp. S55-S62, 2020/04/01/ 2020, doi: https://doi.org/10.1016/j.injury.2020.02.030.

[13] S. A. Bloomfield, "Changes in musculoskeletal structure and function with prolonged bed rest," Med Sci Sports Exerc, vol. 29, no. 2, pp. 197-206, Feb 1997, doi: 10.1097/00005768-199702000-00006.

[14] L. Abdullah, D. E. Davis, P. D. Fabricant, K. Baldwin, and S. Namdari, "Is There Truly "No Significant Difference"? Underpowered Randomized Controlled Trials in the Orthopaedic Literature," J Bone Joint Surg Am, vol. 97, no. 24, pp. 2068-73, Dec 16 2015.

[15] J. G. Wright, "Small Simple Trials: A Strategy for Orthopaedic Randomized Trials," JBJS, vol. 100, no. 14, 2018. [Online]. Available: https://journals.lww.com/jbjsjournal/Fulltext/2018/07180/Small_Simple_Trials__A_Strategy_for_Orthopaedic.13.aspx.

[16] J. C. van der Wouden, A. H. Blankenstein, M. J. H. Huibers, D. A. W. M. van der Windt, W. A. B. Stalman, and A. P. Verhagen, "Survey among 78 studies showed that Lasagna's law holds in Dutch primary care research," Journal of Clinical Epidemiology, vol. 60, no. 8, pp. 819-824, 2007/08/01/ 2007, doi: https://doi.org/10.1016/j.jclinepi.2006.11.010.

[17] H. B. Kitaoka, I. J. Alexander, R. S. Adelaar, J. A. Nunley, M. S. Myerson, and M. Sanders, "Clinical rating systems for the ankle-hindfoot, midfoot, hallux, and lesser toes," Foot Ankle Int, vol. 15, no. 7, pp. 349-53, Jul 1994.

[18] F. D. Naal, F. M. Impellizzeri, and P. F. Rippstein, "Which are the most frequently used outcome instruments in studies on total ankle arthroplasty?," Clin Orthop Relat Res, vol. 468, no. 3, pp. 815-26, Mar 2010.

[19] E. M. Van Lieshout et al., "American Orthopaedic Foot and Ankle Society (AOFAS) Ankle-Hindfoot Score: a study protocol for the translation and validation of the Dutch language version," BMJ Open, vol. 7, no. 2, p. e012884, Feb 27 2017.

[20] E. Pinsker and T. R. Daniels, "AOFAS Position Statement Regarding the Future of the AOFAS Clinical Rating Systems," Foot & Ankle International, vol. 32, no. 9, pp. 841-842, 2011, doi: 10.3113/fai.2011.0841.

[21] N. F. SooHoo, M. Shuler, L. L. Fleming, F. American Orthopaedic, and S. Ankle, "Evaluation of the validity of the AOFAS Clinical Rating Systems by correlation to the SF-36," Foot Ankle Int, vol. 24, no. 1, pp. 50-5, Jan 2003.

[22] F. Pena, J. Agel, and J. C. Coetzee, "Comparison of the MFA to the AOFAS outcome tool in a population undergoing total ankle replacement," Foot Ankle Int, vol. 28, no. 7, pp. 788-93, Jul 2007.

[23] I. N. Sierevelt, L. Beimers, C. J. A. van Bergen, D. Haverkamp, C. B. Terwee, and G. Kerkhoffs, "Validation of the Dutch language version of the Foot and Ankle Outcome Score," Knee Surg Sports Traumatol Arthrosc, vol. 23, no. 8, pp. 2413-2419, Aug 2015.

[24] A. Sharma, D. P. Yadav, H. Garg, M. Kumar, B. Sharma, and D. Koundal, "Bone Cancer Detection Using Feature Extraction Based Machine Learning Model," Comput Math Methods Med, vol. 2021, p. 7433186, 2021.

[25] J. Hirvasniemi, S. Klein, S. Bierma-Zeinstra, M. W. Vernooij, D. Schiphof, and E. H. G. Oei, "A machine learning approach to distinguish between knees without and with osteoarthritis using MRI-based radiomic features from tibial bone," Eur Radiol, vol. 31, no. 11, pp. 8513-8521, Nov 2021.

[26] J. Olczak et al., "Artificial intelligence for analyzing orthopedic trauma radiographs," Acta Orthop, vol. 88, no. 6, pp. 581-586, Dec 2017.

[27] S. W. Chung et al., "Automated detection and classification of the proximal humerus fracture by using deep learning algorithm," Acta Orthop, vol. 89, no. 4, pp. 468-473, Aug 2018.

[28] R. Lindsey et al., "Deep neural network improves fracture detection by clinicians," Proc Natl Acad Sci U S A, vol. 115, no. 45, pp. 11591-11596, Nov 6 2018.

[29] D. H. Kim and T. MacKinnon, "Artificial intelligence in fracture detection: transfer learning from deep convolutional neural networks," Clin Radiol, vol. 73, no. 5, pp. 439-445, May 2018.

[30] T. Urakawa, Y. Tanaka, S. Goto, H. Matsuzawa, K. Watanabe, and N. Endo, "Detecting intertrochanteric hip fractures with orthopedist-level accuracy using a deep convolutional neural network," Skeletal Radiology, vol. 48, no. 2, pp. 239-244, 2019/02/01 2019, doi: 10.1007/s00256-018-3016-3.

[31] P. H. S. Kalmet et al., "Deep learning in fracture detection: a narrative review," Acta Orthop, vol. 91, no. 2, pp. 215-220, Apr 2020.

[32] P. Chea and J. C. Mandell, "Current applications and future directions of deep learning in musculoskeletal radiology," Skeletal Radiology, vol. 49, no. 2, pp. 183-197, 2020/02/01 2020, doi: 10.1007/s00256-019-03284-z.

[33] A. Willems, O. P. van der Jagt, and D. E. Meuffels, "Extracorporeal Shock Wave Treatment for Delayed Union and Nonunion Fractures: A Systematic Review," J Orthop Trauma, vol. 33, no. 2, pp. 97-103, Feb 2019.

[34] M. Kobayashi, R. Chijimatsu, H. Yoshikawa, and K. Yoshida, "Extracorporeal shock wave therapy accelerates endochondral ossification and fracture healing in a rat femur delayed-union model," Biochem Biophys Res Commun, vol. 530, no. 4, pp. 632-637, Oct 1 2020.

[35] C. J. Wang et al., "VEGF modulates angiogenesis and osteogenesis in shockwave-promoted fracture healing in rabbits," J Surg Res, vol. 171, no. 1, pp. 114-9, Nov 2011.

[36] C. J. Wang, F. S. Wang, and K. D. Yang, "Biological effects of extracorporeal shockwave in bone healing: a study in rabbits," Arch Orthop Trauma Surg, vol. 128, no. 8, pp. 879-84, Aug 2008.

[37] O. P. van der Jagt et al., "Unfocused extracorporeal shock wave therapy as potential treatment for osteoporosis," J Orthop Res, vol. 27, no. 11, pp. 1528-33, Nov 2009.

[38] O. P. van der Jagt et al., "Unfocused extracorporeal shock waves induce anabolic effects in osteoporotic rats," J Orthop Res, vol. 31, no. 5, pp. 768-75, May 2013.

[39] D. J. Hak et al., "Delayed union and nonunions: epidemiology, clinical issues, and financial aspects," Injury, vol. 45 Suppl 2, pp. S3-7, Jun 2014.

[40] C. V. Odvina, J. E. Zerwekh, D. S. Rao, N. Maalouf, F. A. Gottschalk, and C. Y. Pak, "Severely suppressed bone turnover: a potential complication of alendronate therapy," J Clin Endocrinol Metab, vol. 90, no. 3, pp. 1294-301, Mar 2005.

[41] A. Giusti, N. A. Hamdy, and S. E. Papapoulos, "Atypical fractures of the femur and bisphosphonate therapy: A systematic review of case/case series studies," Bone, vol. 47, no. 2, pp. 169-80, Aug 2010.

[42] Y. A. Weil, G. Rivkin, O. Safran, M. Liebergall, and A. J. Foldes, "The outcome of surgically treated femur fractures associated with long-term bisphosphonate use," J Trauma, vol. 71, no. 1, pp. 186-90, Jul 2011.

[43] C. W. Oh et al., "Minimally invasive plate osteosynthesis of subtrochanteric femur fractures with a locking plate: a prospective series of 20 fractures," Arch Orthop Trauma Surg, vol. 129, no. 12, pp. 1659-65, Dec 2009.

[44] A. Barquet, L. Francescoli, D. Rienzi, and L. Lopez, "Intertrochanteric-subtrochanteric fractures: treatment with the long Gamma nail," J Orthop Trauma, vol. 14, no. 5, pp. 324-8, Jun-Jul 2000.

[45] B. J. Teo, J. S. Koh, S. K. Goh, M. A. Png, D. T. Chua, and T. S. Howe, "Post-operative outcomes of atypical femoral subtrochanteric fracture in patients on bisphosphonate therapy," Bone Joint J, vol. 96-B, no. 5, pp. 658-64, May 2014.

Chapter 9

Summary

In this chapter the most important findings of this thesis will be presented. The thesis covers three themes that are combined in chapter 7. In the first part of this thesis, we elaborated on surgical treatments for knee and ankle OA. The second part is about objective assessment of bone union. The last part is about the effectiveness of ESWT on bone union. All three themes are introduced in **chapter 1** of this thesis.

Osteoarthritis

In **chapter 2** we compared two surgical techniques to conduct a high tibial osteotomy (HTO) in patients with medial knee OA and varus malalignment. The primary outcome of this randomized controlled trial (RCT) was the difference in accuracy of performed correction between the open wedge HTO and the closed wedge HTO. Accuracy of performed correction was defined as the achieved correction minus the planned correction, based on the hip-knee-ankle angle. The results showed that open wedge HTO slightly under-corrects (0.9°, 95% confidence interval (CI) -2.0 to 0.1) and closed wedge HTO slightly over-corrects (1.0°, 95% CI -0.1 to 2.0) the planned correction. The difference in accuracy of performed correction between the techniques was 1.9 degrees, which was significantly different (95% CI 0.7 to 3.1). No significant differences were found in loss of correction and complications between both groups. It is currently unknown what the long-term effects of under- or over-correction are on OA progression. It seems that both techniques can accurately correct malalignment to the planned correction angle. Therefore, based on the radiological results of this study, both techniques can be used to perform HTO.

In **chapter 3** we investigated the long-term effects of talocrural arthrodesis on the adjacent joints (subtalar, talonavicular and calcaneocuboid joint). Literature has shown that OA is present in adjacent joints after talocrural arthrodesis. However, it is unknown whether OA is present before talocrural arthrodesis, or whether it develops as a consequence of talocrural arthrodesis. A strong feature of our retrospective cohort study was that computed tomography (CT) was available pre-operatively and at follow-up. Our study showed that OA in adjacent joints before talocrural arthrodesis is not increased compared to the contralateral control foot. However, the degree of OA at follow-up was significantly higher compared to pre-operative and also compared to the contralateral foot. It therefore seems that OA develops in adjacent joints after talocrural arthrodesis. However, the degree of OA was not correlated to patient reported outcomes and therefor it seems that the clinical impact of adjacent joint OA is limited.

Bone union assessment

Assessment of bone union is a commonly performed task in clinical practice and studies. It has been shown that progress of consolidation after arthrodesis of foot joints, cannot be accurately assessed from radiographs. Therefore, CT is strongly recommended. However, there is no valid and reliable golden standard method for bone union assessment from CT. In **chapter 4** we conducted a systematic review, which aimed to find a reliable CT parameter that could represent bone union. We were interested in studies that associated CT parameters to actual bone healing, which was assessed by histological or mechanical testing. Studies were included if they were animal studies, created a fracture, assessed bone union with CT, performed mechanical or histological testing, and associated CT-generated outcomes to mechanical or histological testing results. The most common CT parameters that were investigated by the included studies were bone mineral density, bone volume and total callus volume. However, the associations between those parameters and actual bone healing were conflicting, and therefore these parameters do not seem to be reliable for bone union assessment. CT-assessed torsional rigidity and callus density showed the best associations with actual bone union. Therefore, these parameters seem to be the most promising CT parameters to represent bone union.

In **Chapter 5** we presented an overview of currently used methodologies for bone union assessment with CT after foot or ankle arthrodesis. This was done by conducting a systematic review in which studies were included that performed foot or ankle arthrodesis, mentioned radiological of CT follow-up in their abstract and performed postoperative CT in >50% of patients. The most popular method for bone union assessment was by subjectively categorizing bone union into groups (for example: no fusion, partial fusion or total fusion). Although this is a subjective method, the reliability score of this method was acceptable. Another methodology that was frequently used was calculating percentage fusion. Measuring the joint widths on the CT slices does this. Also, all fused parts of the joint should be measured. The sum of the fused parts should be divided by the sum of the joint widths and multiplied by 100, which will result in a percentage fusion. Most studies used a fusion threshold to distinguish between fused and non-fused joints. A fusion threshold of 50% was used most frequently, but this was chosen quite arbitrary. The studies included in our review showed that a fusion threshold of 30% may be more valid to discriminate between fused and non-fused joints. Based on the results of this review we would recommend calculating fusion threshold and applying a 30% fusion threshold for bone union assessment after foot and ankle arthrodesis.

Extracorporeal shock wave therapy (ESWT)

In the last part of this thesis the effectiveness of ESWT on bone union was investigated. In **Chapter 6** we performed a systematic review to present the currently available literature on the effectiveness of ESWT as a treatment for delayed-union and nonunion fractures. In this review, studies were included that treated patients with delayed-union or nonunion fractures with ESWT. The systematic review showed that the average union rate after ESWT in delayed-unions was 86% and in nonunions 75%. No severe complications occurred after ESWT. Patients with nonunions that were treated surgically showed union rates of 81%. After surgery, several studies reported major complications like hardware removal or wound infections. Based on this study it seems that ESWT is as effective as surgical treatment for treatment of delayed-unions and nonunions. Also, it seems that ESWT is safer than surgical treatment. However, generally the quality of the included studies was low as most studies were cohort studies. Therefore, more high-level evidence studies should be done to prove the effectiveness of ESWT for delayed-union and nonunion treatment.

In **chapter 7** we assessed whether ESWT can reduce the number of delayed-unions after talocrural arthrodesis. In this RCT, patients were randomized between per-operative ESWT (intervention group) or per-operative sham-ESWT (control group). The primary outcome of this RCT was the number of delayed unions assessed on CT twelve weeks after talocrural arthrodesis. Union was assessed as recommended in chapter 5, by calculating the fusion percentage and applying a fusion threshold of 30%. The results showed that the number of delayed unions in the intervention group was 3 out of 19 (16%), and in the control group 6 out of 22 (27%). The number of delayed unions was not significantly different between the groups ((Risk difference=0.11; 95% CI 0.43 – 9.42; p=0.38). Although the number of delayed unions seem to be lower in the intervention group, the difference was not significant. However the promising results of this study, which tend to a favorable effect of ESWT, should encourage the conduction of more high quality RCT.

Appendices

Nederlandse samenvatting

Dankwoord

Curriculum Vitae

PhD portfolio summary

List of publications

Nederlandse samenvatting

Dit proefschrift omvat drie thema's die uiteindelijk samenkomen in hoofdstuk 7 van dit proefschrift. De thema's in dit proefschrift zijn 1) operatieve behandeling van artrose, 2) objectieve bepaling van botdoorbouw en 3) het effect van extracorporale shock wave therapie (ESWT) op botdoorbouw. Alle drie de thema's zijn in **hoofdstuk 1** van dit proefschrift geïntroduceerd. In hoofdstuk 7 werd de effectiviteit van ESWT op botdoorbouw onderzocht, in patiënten die als gevolg van artrose in het talocrurale gewricht een arthrodese ondergingen. Hiervoor is het belangrijk dat alle drie de thema's aan bod komen.

Operatieve behandeling van artrose

Artrose is een multifactoriële gewrichtsaandoening waarbij niet alleen het kraakbeen is beschadigd, maar ook andere weefsel in het gewricht zijn aangedaan. Patiënten met artrose ervaren vaak klachten als pijn, stijfheid, instabiliteit en zwelling. Bij beginnende artrose worden vaak eerst conservatieve therapieën ingezet, zoals fysiotherapie voor het verbeteren van spierkracht, aerobe fitheid en het reduceren van gewicht. Ook kunnen symptomen bestreden worden door het gebruik van pijnstilling. Indien conservatieve therapieën niet (meer) effectief zijn, kan een operatieve ingreep overwogen worden. Welke operatieve ingreep nodig is hangt af van de ernst van de artrose en welk gewricht is aangedaan. In **hoofdstuk 2** van dit proefschrift hebben wij onderzoek gedaan naar een hoge tibiakop osteotomie (HTO). Dit is een operatie die kan worden uitgevoerd bij patiënten met matige tot ernstige mediale gonartrose (artrose aan de binnenkant van de knie) en een varus beenstand (o-stand). Tijdens een HTO wordt de varus beenstand gecorrigeerd naar een lichte valgusstand (x-stand). Hierdoor wordt het mediale compartiment van de knie ontlast en nemen de pijnklachten af. Voor een goed resultaat is het belangrijk dat de correctie nauwkeurig wordt uitgevoerd. Zowel een over- als onder-correctie zouden op termijn negatieve gevolgen kunnen hebben.

Een HTO kan middels verschillende technieken worden uitgevoerd. In hoofdstuk 2 hebben wij twee technieken met elkaar vergeleken, namelijk de open wig techniek en de gesloten wig techniek. Het doel van deze gerandomiseerde studie was om uit te zoeken of er een verschil is in de nauwkeurigheid van de correctie tussen de technieken. De nauwkeurigheid van de correctie werd bepaald door het verschil tussen de vooraf bepaalde correctie en de daadwerkelijk behaalde correctie te berekenen. De resultaten van de studie lieten zien dat de open wig HTO iets onder-corrigeert (0.9°, 95% betrouwbaarheidsinterval (BI) -2.0 tot 0.1), terwijl de gesloten wig HTO iets over-corrigeert (1.0°, 95% BI -0.1 tot 2.0). Het verschil in nauwkeurigheid tussen de twee technieken was

1.9°, wat een significant verschil was (95% BI 0.7 tot 3.1). Ondanks dat de nauwkeurigheid van correctie significant verschillend is tussen de technieken, lijken beide technieken een nauwkeurige correctie te behalen. Omdat op basis van de literatuur niet bekend is wat de lange termijn effecten van een onder- of over-correctie zijn, lijken beide technieken op basis van deze radiologische resultaten geschikt voor het uitvoeren van een HTO.

In **hoofdstuk 3** van dit proefschrift hebben we onderzoek gedaan naar een operatieve ingreep voor patiënten met talocrurale artrose (artrose in het bovenste spronggewricht). Bij patiënten met ernstige talocrurale artrose kan men een arthrodese uitvoeren. Bij deze ingreep wordt het resterende kraakbeen uit het talocrurale gewricht verwijderd en wordt het talocrurale gewricht vastgezet. Vervolgens groeien de botuiteinden in het talorurale gewicht aan elkaar vast, waardoor er geen beweging meer mogelijk is in het gewricht en de pijnklachten afnemen. Uit de literatuur blijkt dat na een talocrurale arthrodese, er vaak artrose voorkomt in de omliggende gewrichten (talonaviculaire, subtalaire en calcaneocuboid gewrichten). Dit komt mogelijk doordat deze gewrichten zwaarder worden belast na een arthrodese, omdat zij moeten compenseren voor het vastgezette talocrurale gewricht. Op basis van de huidige literatuur is het echter onduidelijk of de artrose in de omliggende gewrichten ontstaat als gevolg van de arthrodese, of dat de artrose al aanwezig was voor de arthrodese. Om dit verder te onderzoeken hebben wij een retrospectief cohort onderzoek uitgevoerd. Om de mate van artrose nauwkeurig te kunnen beoordelen hebben wij van alle deelnemende patiënten een preoperatieve computer tomografie scan (CT) en een bilaterale (van beide voeten/enkels) follow-up CT. Hierbij werd de CT van de niet geopereerde kant gebruikt als gezonde controle scan. Onze resultaten lieten zien dat de mate van artrose in de omliggende gewrichten vóór de arthrodese niet verschillend was van de gezonde controle kant. Voor de arthrodese was er dus geen sprake van meer artrose in de omliggende gewrichten. De mate van artrose in de omliggende gewrichten was wel significant verschillend tussen de preoperatieve CT and follow-up CT, wat dus duidt op een toename van artrose in de omliggende gewrichten na de arthrodese. Ook was er een significant verschil in mate van artrose tussen de arthrodese kant en gezonde controle kant. Hieruit blijkt dat er meer artrose is in de gewrichten van de geopereerde voet/enkel dan aan de gezonde kant. Op basis van dit onderzoek lijkt het dus dat de artrose in omliggende gewrichten ontstaat als gevolg van de arthrodese. Overigens vonden wij geen verbanden tussen de mate van artrose in de omliggende gewrichten en patiënt gerapporteerde uitkomsten. Het lijkt er dus op dat de artrose in de omliggende gewrichten weinig klachten veroorzaakt bij de patiënt.

Objectieve bepaling van botdoorbouw

Het beoordelen van de botdoorbouw na een breuk, arthrodese of osteotomie is dagelijkse praktijk voor orthopeden, radiologen en sportartsen. Eerder onderzoek heeft aangetoond dat de botdoorbouw na een arthrodese in de voet niet nauwkeurig beoordeeld kan worden op basis van röntgenfoto's. Het is daarom beter om de botdoorbouw na een arthrodese te beoordelen op basis van CT. Echter, er bestaat geen gouden standaard methode om botdoorbouw objectief en nauwkeurig te bepalen aan de hand van CT. Daarom hebben wij in **hoofdstuk 4** onderzocht welke CT parameter het meest geschikt is om botdoorbouw mee te bepalen. Hiervoor hebben wij een systematische review uitgevoerd. Voor deze systematische review hebben wij studies geïncludeerd die CT parameters correleerden aan daadwerkelijke botdoorbouw. De mate van daadwerkelijke botdoorbouw werd gepaald aan de hand van histologische of mechanische testen. Omdat dit niet in patiënten kan worden uitgevoerd, hebben wij ons gericht op onderzoeken met proefdieren. In deze onderzoeken werd bij een proefdier een breuk gecreëerd, waarbij na enig weken een CT werd gemaakt en direct daarna een histologische of mechanische test werd uitgevoerd. Vervolgens keken de studies of er een correlatie was tussen de CT parameters en de uitkomsten van de histologische of mechanische testen. De meest onderzochte CT parameters waren botdichtheid, bot volume en totale callus volume. Echter, de in de studies gerapporteerde correlaties tussen deze parameters en de daadwerkelijke botdoorbouw waren erg verschillend en daarom lijken deze CT parameters geen goede maat om de botdoorbouw mee te bepalen. Voor twee CT parameters werden wel consequent sterke correlaties gevonden met de daadwerkelijke botdoorbouw. Dit waren CT-bepaalde torsiestijfheid en callus dichtheid. Deze CT parameters lijken daarom het meest geschikt om botdoorbouw mee te kunnen bepalen. Er zal echter nog meer onderzoek gedaan moeten worden om een betrouwbare methode te ontwikkelen op basis van deze CT parameters.

Omdat er geen gouden standaard is voor de bepaling van botdoorbouw met CT na een voet of enkel arthrodese, worden er op dit moment verschillende methodes gebruikt. In **hoofdstuk 5** is een overzicht gegeven van de op dit op dit moment gebruikte methodes en is ook gekeken naar de betrouwbaarheid van deze methodes. Uit deze systematische review blijkt dat een veelgebruikte methode het subjectief categoriseren van de mate van botdoorbouw is. De CT beoordelaar kan dan bijvoorbeeld kiezen uit geen botdoorbouw, matige botdoorbouw of volledige botdoorbouw. De betrouwbaarheid van deze methode bleek acceptabel. Een andere methode is het berekenen van het percentage botdoorbouw. Hiervoor wordt op meerder CT coupes de breedte van de gewricht en de breedte van het gefuseerde deel gemeten. Vervolgens worden alle gewrichtsbreedtes bij elkaar opgeteld en ook alle gefuseerde delen. Hiermee wordt dan het percentage botdoorbouw berekend. De meeste studies die deze methode hanteerden gebruikten

vervolgens een drempelwaarde waarbij men stelt dat vanaf 50% botdoorbouw een gewricht gefuseerd is. Deze methode resulteerde in een uitstekende betrouwbaarheid. Echter, de drempelwaarde was erg arbitrair gekozen. Verschillende studies tonen aan dat vanaf 30% botdoorbouw een gewricht als gefuseerd mag worden beschouwd. Op basis van dit onderzoek adviseren wij om de botdoorbouw na een voet of enkel arthrodese te bepalen op een CT door het percentage botdoorbouw te berekenen en een drempelwaarde van 30% aan te houden om onderscheid te maken tussen gefuseerde en niet gefuseerde gewrichten.

Extracorporale schok wave therapie

In het laatste deel van dit proefschrift is de effectiviteit van ESWT op botdoorbouw onderzocht. ESWT, ook wel drukgolftherapie genoemd, is een niet invasieve behandelmethode die mogelijk botdoorbouw zou kunnen stimuleren. Op dit moment wordt ESWT met name ingezet bij patiënten met vertraagd helend of niet helende breuken. **Hoofdstuk 6** omvat een systematische review die de op dit moment beschikbare literatuur toont met betrekking tot de effectiviteit van ESWT op vertraagd helende en niet helende breuken. In deze systematische review werden studies geïncludeerd die vertraagd helende of niet helende breuken behandelden met ESWT. De resultaten van de review lieten zien dat na de ESWT 86% van de vertraagd helend botbreuken alsnog consolideerde. In de niet helende botbreuken was dit 75%. Op dit moment worden niet-helende botbreuken vaak operatief behandeld. De review liet zien dat na een operatie 81% van de niet helende botbreuken consolideert. Uit de studies bleek dat ESWT geen ernstige bijwerkingen veroorzaakt, terwijl na een operatieve behandeling wondinfecties optraden en soms een tweede operatie nodig was om het osteosynthese materiaal te verwijderen. Op basis van dit onderzoek lijkt het dus alsof ESWT net zo effectief is als de operatieve behandeling van niet genezende breuken. Daarnaast is bij ESWT minder kans op ernstige complicaties. Echter, de meeste studies die in dit onderzoek werden geïncludeerd waren van slechte kwaliteit. Er zal dus meer kwalitatief goed onderzoek uitgevoerd moeten worden om de effectiviteit van ESWT te bewijzen.

In **hoofdstuk 7** van dat in dit proefschrift, komen alle eerder besproken thema's samen. In dit gerandomiseerde onderzoek is gekeken of ESWT het aantal vertraagd genezende arthrodeses kan verlagen. In dit onderzoek werden patiënten geïncludeerd die een talocrurale arthrodese ondergingen als gevolg van ernstige talocrural artrose. Patiënten werden gerandomiseerd tussen peroperatieve ESWT (interventie groep) of peroperatieve placebo ESWT (controle groep). De botdoorbouw na de arthrodese werd beoordeeld aan de hand van de in hoofdstuk 5 geadviseerde methode. Hierbij werd het percentage botdoorbouw berekend op basis van de CT die gemaakt was 12 weken na de

arthrodese. Bij minder dan 30% botdoorbouw werd een arthrodese geclassificeerd als vertraagd genezend. De resultaten van dit onderzoek lieten zien dat het aantal vertraagd genezende arthrodeses in de interventie groep 3 van de 16 (19%) was, waarbij dit in de controlegroep 6 van de 22 (27%) was. Het aantal vertraagd genezende arthrodeses bleek niet significant verschillend tussen de groepen (95% CI 0.43 – 9.42; p=0.38). Ondanks dat er geen significant verschil in het aantal vertraagd genezende arthrodeses werd gevonden, lijkt er op basis van de absolute aantallen wel sprake van een afname in vertraagd genezend arthrodeses. Op basis van deze veelbelovende trend, zou het daarom te adviseren zijn om in de toekomst meer onderzoek te doen naar de effectiviteit van ESWT.

Dankwoord

Dit proefschrift was niet tot stand gekomen zonder de hulp, adviezen en steun van veel mensen om mij heen. Graag wil ik iedereen die heeft bijgedragen aan dit proefschrift enorm bedanken. Een aantal personen wil ik in het bijzonder bedanken.

Beste Sita, ik was al even bezig met de SPARKLE-studie, toen jij officieel tot mijn promotor werd benoemd. Van ons eerste officiële promotieoverleg herinner ik me vooral dat jij heel doortastend was, en er duidelijke afspraken werden gemaakt, de koers voor mijn promotie werd uitgezet! Met een duidelijke koers en doel werd het hele traject voor mij een stuk overzichtelijker. Ik wil je enorm bedanken voor jouw kritische blik, expertise en efficiënte overleggen.

Beste Duncan, ook jij bent enorm belangrijk geweest voor het tot stand komen van dit proefschrift. De samenwerking met jou heb ik als heel erg prettig ervaren. Je was altijd bereid om even mee te kijken, en mij te voorzien van nieuwe ideeën, adviezen en stof tot nadenken. Naast jouw wetenschappelijke en vakkundige inbreng toonde je ook oprechte interesse in mij en kwam je vaak even binnenwandelen in HS-105 en later NC-424. Je bent je steeds voor mij blijven inzetten en in mijn proefschrift blijven geloven. Hiervoor wil ik je heel erg bedanken, dankzij jou kreeg ik de tijd om dit proefschrift te voltooien.

Beste Olav, iets minder aanwezig in het dagelijks reilen en zijlen maar vanaf een afstand heb jij het hele promotietraject altijd gevolgd en gesteund. Allereerst wil ik je bedanken voor het bedenken en opzetten van de SPARKLE-studie. Zonder deze studie was dit proefschrift niet tot stand gekomen. Ik vond het een ontzettend leuk onderzoek om uit te voeren, waarbij ook nog eens een grote klinische relevantie was. Daarnaast was jij altijd bereikbaar voor een overleg. Jouw frisse blik zorgde bij mij vaak voor nieuwe inzichten, dank daarvoor!

Beste Eline, dankjewel voor alle gezelligheid in de afgelopen jaren. Een goede werksfeer draagt enorm bij aan het werkplezier, en hierin ben jij een belangrijke factor geweest. Daarnaast was je ook altijd bereid om mee te denken en had je vaak goede ideeën. Ik heb met heel veel plezier met jou gewerkt en ben heel blij dat je mijn paranimf wilt zijn.

Beste Suzie, wie had in 2008 kunnen bedenken, dat wij nu hier zouden staan. De afgelopen jaren hebben onze wegen zich meermaals gekruist. Wat begon als studiegenoten, veranderde al snel in een vriendschap en huisgenoten, collega's, buren en nu paranimf. Dankjewel dat je mij op deze belangrijke dag wilt bijstaan!

Beste Max, ondanks dat jij geen deel uit maakt van mijn promotieteam was jij wel altijd bereid om even mee te denken en kon ik altijd bij jou terecht met vragen. Naast jouw wetenschappelijke kunde en kennis zorgde jij ook voor een fijne sfeer waarin ik me erg heb thuis gevoeld. Dankjewel voor alle wetenchappelijke overleggen en gezellige momenten. Mede hierdoor kijk ik terug op een hele leerzame en gezellige periode. Ik ben dan ook blij dat we op het eind van mijn promotietraject toch nog 'officieel' hebben kunnen samenwerken aan de HTO-studie.

Beste Edwin, jij was altijd bereid en enthousiast om met ons mee te denken over de radiologische vraagstukken in dit proefschrift. Heel erg bedankt hiervoor!

Beste Menno, René, Jarno, Floris en Wouter, dank voor jullie bijdrage aan de SPARKLE studie vanuit het Maasstad ziekenhuis. Menno, dankzij jou konden we de SPARKLE studie starten in het Maasstad ziekenhuis wat leidde tot een flinke toename in de inclusiesnelheid, bedankt hiervoor! René, Jarno en Floris, super fijn dat jullie na het vertrek van Menno bereid waren om de SPARKLE studie voort te zetten! Wouter, dank dat jij de rol van hoofdonderzoeker op je wilde nemen, dit was van groot belang voor de continuering van de SPARKLE studie in het Maasstad!

Beste Inke en Annette, in de overvolle operatieplanning kregen jullie het toch steeds weer voor elkaar om plekjes vrij te maken voor de SPAKRLE patiënten, bedankt!

Beste Simone en Annet, dank dat ik altijd bij jullie terecht kon met de meest willekeurige vragen. Jullie enorme vriendelijkheid en behulpzaamheid vormt wat mij betreft een belangrijke pijler voor de fijne werksfeer binnen de afdeling!

Beste kamergenoten van NC-424 en HS-105 (Suus, Mark, Joost, Erwin, Arco, Lichelle, Abi, Pim, Iris, Britt, Sabine, Floris, Michiel, Jinchi, Delong, Erin, Fleur, Joshua, Noor en Rintje), mede door jullie ging ik elke dag met plezier naar het EMC. De sfeer in de onderzoekskamer was altijd goed en we hebben heel wat afgelachen. Naast een hoop gezelligheid was er ook ruimte voor serieuze gesprekken over onze onderzoeken en problemen waar we tegenaan liepen. De jaarlijkse onderzoeksuitjes waren altijd een succes met als hoogtepunt voor mij toch wel ons rondje IJsselmeer op de fiets! Bedankt allemaal voor jullie gezelligheid, support en mooie herinneringen!

Beste Caner, Mauro, Christa en Sebastian, binnen verschillende onderzoeken van dit proefschrift hebben jullie mij enorm geholpen met het uitvoeren van metingen, screenen van patiënten of screenen van heel veel artikelen. Jullie bleven steeds enthousiast en

betrokken bij de lopende projecten. Heel erg bedankt voor jullie inzet, betrokkenheid en doorzettingsvermogen, zonder jullie was dit proefschrift nog lang niet klaar geweest!

Beste Hanneke, Jeroen, Joost en Fleur, naast mijn promotie mocht ik de afgelopen jaren bijdragen aan het uitvoeren van de FIXIT studie. Ik heb dit met heel erg veel plezier gedaan, mede door de fijne samenwerking met jullie! Bedankt hiervoor!

Lieve Chris, Iris, Mark, Luuk, Arjan en Robin, sinds de middelbare school zijn wij elkaar niet meer uit het oog verloren. Dank voor jullie support tijdens mijn promotie en de bijzondere vriendschappen!

Lieve Wendy, Anne, Saskia en Paulien, tijdens onze studie hebben wij elkaar ontmoet en sindsdien hebben we al veel samen meegemaakt. Bedankt dat jullie er altijd voor mij waren tijdens mijn promotietraject! Ik hoop dat we in de toekomst nog maar heel veel mooie momenten samen mogen delen! #WASSAP4ever!

Lieve Angela, Willem, Linda, Ewoud, Björn en Mette, een leukere schoonfamilie had ik mij niet kunnen wensen. Tijdens mijn promotietraject leerde ik Martijn kennen en werd ik aan jullie geïntroduceerd tijdens een uitje naar de Efteling. Ik had toen niet kunnen vermoeden hoe erg ik me thuis zou gaan voelen in de Rödel familie. Bedankt dat jullie er altijd voor mij (en ons) zijn!

Lieve Mark, Isa, Carmen en Lola, bedankt dat jullie er zijn! Mark, sinds kleins af aan hebben wij altijd een goede band gehad, en ben jij voor mij een geweldige broer geweest. Ik vind het heel bijzonder dat wij nu beiden een prachtig gezin hebben en deze bijzondere levensfase samen kunnen meemaken. Isa, Carmen en Lola, ustedes son increíbles!

Lieve Pap en Mam, jullie zijn enorm belangrijk voor mij! De liefdevolle, warme en veilige thuishaven waarin ik ben opgegroeid heeft mij gebracht tot waar ik nu ben. Dankzij jullie onuitputtelijke liefde en steun, had ik het vertrouwen om dit promotietraject aan te gaan. Bedankt dat ik altijd op jullie terug kan vallen!

Lieve Martijn, Liza en Jonas, jullie zijn het allermooiste dat mij is overkomen. Martijn, dankjewel voor alle steun en liefde die jij mij in de afgelopen jaren hebt gegeven. Ik houd heel erg veel van jou!

Curriculum vitae

Judith Annika Willems werd geboren op 26 september 1990 te Rotterdam. Tussen 2002 en 2008 doorliep zij het VWO op het Sint Laurenscollege te Rotterdam. Hierna begon zij aan de studie Bewegings-wetenschappen aan de Vrije Universiteit te Amsterdam, waar zij in 2012 haar bachelor diploma behaalde. Het master diploma bewegings-wetenschappen volgde in 2014. In het kader van de Master opleiding werd een onderzoeksstage gedaan bij het Peter Harrison Centre for Disability Sport aan de Loughborough University te Engeland.

Na het behalen van haar master diploma bleef ze nog enige tijd werken als onderzoeker op de afdeling geriatrie in het Amsterdam UMC. In juli 2015 werd ze in deeltijd aangenomen als promovendus bij de afdeling Orthopedie en Sportgeneeskunde van het Erasmus MC. Het promotietraject bestond primair uit het uitvoeren van een gerandomiseerde studie naar het effect van drukgolftherapie op de botdoorbouw na een enkelartrodese. Deze studie werd uitgevoerd in samenwerking met de afdeling Orthopedie van het Maasstad ziekenhuis. Al snel na de start van het promotietraject kon ze haar promotietraject combineren met het bieden van ondersteuning bij een andere gerandomiseerde studie op de afdeling Orthopedie en Sportgeneeskunde. Tijdens het laatste jaar van haar promotietraject was ze tevens werkzaam als onderzoeker bij de afdeling Orthopedie in het Rijnstate ziekenhuis te Arnhem.

Annika Willems is sinds 2021 getrouwd met Martijn Rödel. Samen met hun kinderen Liza (2021) en Jonas (2023) wonen zij in Huissen. Annika is momenteel werkzaam als beleidsadviseur onderzoek aan de Hogeschool van Arnhem en Nijmegen bij de academie sport & bewegen en academie paramedische studies.

PhD portfolio

Personal details	
Name	Judith Annika Willems
Department	Department of Orthopaedics and Sports Medicine
PhD period	2015 - 2023
Promotor	Prof. dr. S.M.A. Bierma-Zeinstra
Supervisors	Dr. D.E. Meuffels
	Dr. O.P. van der Jagt

	Year	**ECTS**
General Courses		
- Erasmus MC Research Inegrity	2015	0.3
- Basic course LimeSurvey and GemsTracker	2016	1.5
- Erasmus MC Systematic Literature Search in Pubmed and other databases	2016	0.75
- Erasmus MC EndNote	2016	0.2
- Erasmus MC BROK (Basic course Rules and Organisation for Clinical Researchers)	2016	1.5
- Erasmus MC Biomedical English writing and Communication	2016	1.5
- Nihes Advanced Clinical Trials (EWP10)	2017	1.9
- Nihes Courses for the Quantitative Researcher	2017	1.4
- BROK re-registration	2020	0.75
- Presentation workshop	2020	1.5
- Nihes Biostatistical Methods I: Basic Principles (CC02)	2020	5.7
- Nihes Logistic Regrssion (ESP66)	2022	1.4
Specific courses		
- ISMST instructional certification course	2017	1.0
Oral and poster presentations		
- ISMST congress, San Sebastian, Spain *Systematic review on the effectiveness of extracorporeal shock wave treatment (ESWT) as an alternative treatment for nonunions*	2017	1.0
- NOV congress, Leeuwarden, The Netherlands *Open of gesloten hoge tibiakop osteotomie? Een randomised controlled trial*	2022	1.0

- OVO congress, Den Bosch, The Netherlands *SPARKLE-studie*	2019	1.0
- NOV congress, Leeuwarden, The Netherlands *Shockwave therapie bij talocrurale arthrodeses: een randomised controlled trial (SPARKLE-studie)*	2022	1.0

Conferences		
- NOV Najaarscongres, Leeuwarden, The Netherlands	2021	0.5
- NOV Najaarscongres, Rotterdam, The Netherlands	2018	0.5
- National Shock wave congress, Zwolle, The Netherlands	2019	0.5
- NOV Najaarscongres, Arnhem, The Netherlands	2019	0.5

Teaching activities and student supervision		
- Teaching master students during their research internship	2016 - 2021	2.5
- Supervising workshop bone pathology bachelor students	2017, 2018 & 2019	2.0
- Supervising writing systematic review bachelor students	2017 & 2019	1.0
- Supervising writing assignment minor Orthopedic sports traumatology	2019 & 2020	1.5
- Presenting Showcase for the minor Orthopedic sports traumatology	2020	0.5

Total ECTS		34.6

List of publications

Willems A, van der Jagt OP, Meuffels Reig SE, Baart S, Bloembergen M, Bierma-Zeinstra SMA, Meuffels DE. Extracorporeal shock wave therapy (ESWT) for talocrural arthrodesis; a double-blinded randomized controlled trial. Submitted

Willems A, Duivenvoorden T, van Es EM, Meuffels DE, Bierma-Zeinstra SMA, Bos PK, Reijman M. Accuracy of performed correction of open versus closed wedge high tibial osteotomies, with locking plate fixation: A randomized controlled trial. Submitted

Willems A, Minnaard M, Oei EHG, Bierma-Zeinstra SMA, Meuffels DE. Talocrural arthrodesis increases osteoarthritis severity in adjacent joints; a midterm computed tomography follow-up study. Submitted

Willems A, Houkes CM, Bierma-Zeinstra SMA, Meuffels DE. How to assess consolidation after foot and ankle arthrodesis with computed tomography. A systematic review. Eur J Radiol. 2022 Sep 5;156:110511.

Willems A, Içli C, Waarsing JH, Bierma-Zeinstra SMA, Meuffels DE. Bone Union Assessment with Computed Tomography (CT) and Statistical Associations with Mechanical or Histological Testing: A Systematic Review of Animal Studies. Calcif Tissue Int. 2022 Feb;110(2):147-161.

Willems A, van der Jagt OP, Meuffels DE. Extracorporeal Shock Wave Treatment for Delayed Union and Nonunion Fractures: A Systematic Review. J Orthop Trauma. 2019 Feb;33(2):97-103.

Willems A, Paulson TA, Keil M, Brooke-Wavell K, Goosey-Tolfrey VL. Dual-Energy X-Ray Absorptiometry, Skinfold Thickness, and Waist Circumference for Assessing Body Composition in Ambulant and Non-Ambulant Wheelchair Games Players. Front Physiol. 2015 Nov 27;6:356.