MRI FOR THE INITIAL EVALUATION OF ACUTE WRIST, KNEE, AND ANKLE TRAUMA

Jeroen J. Nikken

The work of this thesis was conducted at the department of Radiology and the department of Epidemiology & Biostatistics of the Erasmus MC Rotterdam, the Netherlands.
The study was supported by the department of Radiology, the Revolving Fund from the Erasmus University Medical Center Rotterdam and by an unrestricted grant from Esaote S.p.A., Genoa, Italy.
Financial support by the department of Radiology of the Erasmus MC and by Schering Nederland BV for the printing of this thesis is gratefully acknowledged.

ISBN 90-9017191-6

Lay-out : Andries W. Zwamborn
Cover
photography : Jeroen Stout
design : Andries.W. Zwamborn
after an idea of Hugh Turvey

Printed by Ridderprint B.V. Ridderkerk

MRI FOR THE INITIAL EVALUATION OF ACUTE WRIST, KNEE, AND ANKLE TRAUMA

MRI voor de initiële evaluatie van acuut pols, knie en enkel letsel

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
Rector Magnificus

Prof.dr.ir. J.H. van Bemmel

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
woensdag 1 oktober 2003 om 9.45 uur

door

Jeroen Joop Nikken
geboren te Den Haag.

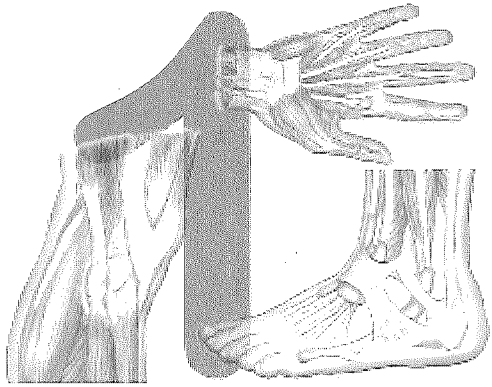
PROMOTIECOMMISSIE

Promotoren: Prof.dr. M.G.M. Hunink
Prof.dr. G.P. Krestin

Overige leden: Prof.dr. B.W. Koes
Prof.dr. J.A.N. Verhaar
Prof.dr. A.B. van Vugt

Copromotor: Dr. A.Z. Ginai

Contents



Introduction

Magnetic resonance imaging has obtained a firm position in the diagnostic evaluation of musculoskeletal abnormalities. It is, however, not widely used as a routine examination tool in the initial examination of acute trauma of wrist, knee and ankle. The high costs and the long duration of the examination are major hindrances for the use of MRI in this respect. Moreover, whether MRI in acute joint trauma is useful is a matter of debate. The issue about the yet largely unknown potential of MRI in acute skeletal trauma has recently been brought up by Rogers, the editor in chief of the American Journal of Radiology, stating that there is still a lot to know and much to be learned about the potential role of MR imaging in acute skeletal trauma (1, 2). Reports about the effect of MRI in acute knee injury on treatment decisions are conflicting; whereas Maurer found that MRI reduced the number of arthroscopic procedures, improved clinician diagnostic certainty, and assisted management decisions (3), Odgaard found in 90 consecutive patients that the MRI result changed treatment in only 6 cases and did not prevent any planned arthroscopy (4). For the wrist the effect of early MRI after trauma has only been studied in relation to scaphoid fractures. These studies show that MRI is very sensitive in the detection of fractures of the scaphoid bone that are not evident on plain radiographs (5-7), but no formal cost-effectiveness studies on this subject from the societal perspective have been published. The effect of MRI on the management of acute ankle injury has only been studied in children with suspected physeal ankle fractures. Carey found that MRI changed management in 5 of 14 children with suspected or known growthplate injury (8), however, Lohman found no change in management after MRI examination in 60 consecutive children with suspected ankle fracture or ankle ligament tear (9).

In this thesis we study the application of MRI in acute trauma of wrist, knee, and ankle, evaluating its potentials, its effects, and its costs. Our aim was to use MRI in all patients with acute trauma of wrist, knee, and ankle, without increasing the overall costs to society, potentially even reducing these costs.

We used a relatively inexpensive low field dedicated extremity MRI system (Artoscan M, Esaote, Genoa, Italy) for our study. There has been debate regarding the diagnostic performance of low field MRI versus high field MRI (10-25). However, most studies conclude that the diagnostic quality of low field MRI approaches the high field systems. We feel that skepticism toward low field MRI is often based on first impressions of the low field MRI images and not on data of comparative clinical trials.

In chapter 2 we present a meta-analysis of earlier published studies on the diagnostic performance of MRI of the menisci and cruciate ligaments combined with an assessment of the effect of magnetic field strength on diagnostic performance.
We developed a shortened scanning protocol on the low field MRI system, reducing both time and costs of the examination to be used for initial evaluation of wrist, knee, and ankle injury (Chapter 3). A reduction in examination time was necessary since a standard MRI examination takes about 30 to 45 minutes, which would be too long and probably too expensive for routine examination of all patients with acute joint trauma.
In Chapter 4, 5, and 6 we describe a multivariable logistic regression analysis for acute wrist, knee, and ankle trauma to assess if the addition of a short MRI examination contributes in discriminating between patients that will need additional therapy after their initial visit to the hospital and patients that can be sent home without need for follow-up. Furthermore we study if the short MRI examination can replace radiography for this purpose.

The strategy of performing a short MRI examination in addition to the current diagnostic work-up is only feasible if it is cost-effective. However, it is not very likely that an initial short MRI examination would lead to an overall increase in quality of life in the long term if all patients with acute trauma of wrist, knee, or ankle were to undergo a short MRI examination. It may be true for the radiographically occult scaphoid fracture, since a delay in treatment could lead to pseudo-arthrosis, osteonecrosis and arthrosis in the end. However, for the majority of possible lesions a delay in diagnosis and treatment is not necessarily detrimental and therefore we did not anticipate a general increase in long term quality of life as a result of the MRI examination. We rather expected the advantage to be a shorter convalescence in a subgroup of patients in whom the earlier final diagnosis would lead to earlier treatment and an earlier recovery. An earlier recovery is favorable for the patient and may also reduce the time off work and thus save costs. For the radiology department an extra MRI examination would always be more expensive. For the health care payer (e.g. insurance company) there may be some financial advantages: potentially less follow-up visits, less temporary treatment measures if the diagnosis is not yet clear at initial evaluation, and possibly less diagnostic procedures during follow-up, including (diagnostic) arthroscopies. There is also the possibility of a false positive MRI result, which could lead to unwarranted treatment, which would increase the costs and decrease patient utility. On the other hand, cost-savings can be expected from the societal perspective by a reduction in lost productivity.

The associated cost reduction may have a major impact on the overall costs (26) and may well be the decisive factor in the cost-effectiveness analysis. In Chapter 7 both the costs and the effects of performing a short MRI examination in addition to history taking, physical examination and radiography are assessed from the societal perspective for acute trauma of the wrist, knee, and ankle.
Chapter 8 focuses on improving the cost-effectiveness of MRI in patients with acute knee injury by identifying a subgroup of patients that would profit most from the early MRI examination.

Socio-economic background

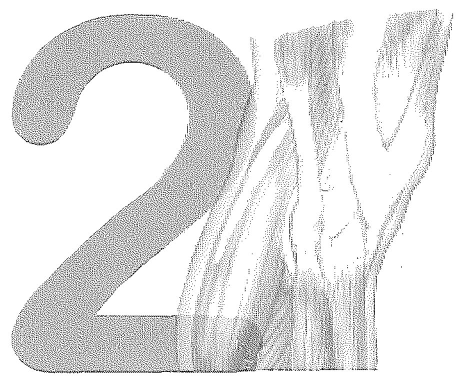
The underlying question in this study from a socio-economic point of view is how we should (re)allocate our financial and medical resources. The theoretical basis for such an analysis is founded in the welfare theory, based on the utilitarian principle of maximum social benefit (27). In the welfare theory several criteria are used: the Pareto criterion for efficiency (introduced in 1890 by Vilfredo Pareto, an Italian sociologist and economist) implies that certain allocation of means is efficient if nobody in that situation can gain anymore without a reduced gain for somebody else. A state is Pareto Efficient if no one's utility can be increased without decreasing someone else's utility. Since this criterion almost certainly prevents any improvement for society as a whole, the Kaldor Hicks criterion was introduced: a certain allocation of means should be implemented if those who experience a reduced gain by this allocation could be compensated. A more lenient point of view is the Social Utility Maximization: under this criterion a certain allocation of means is efficient if it maximizes the total utility of society, even if some people are made worse off by the allocation (28). In the medical setting these principles form the basis of modern cost-effectiveness analysis. For the current study the question would be, considering all costs, benefits, and drawbacks, would society be willing to pay for an additional MRI examination in all patients with acute wrist, knee, and ankle trauma?

The relevance of this study is proportional to the number of wrist, knee, and ankle injuries that occur annually and the resulting time-from-work, time-from-usual-activities, and the costs of medical management. To give an idea of the order of magnitude of sports injuries: over 500 million Euro were spent on the management of accidental falls (excluding motor vehicle accidents) in the Netherlands in the year 1994 (Polder ea. Kosten van ziekten in Nederland in 1994). Ankle, knee, and wrist trauma represent a substantial proportion of such accidents.

References

1. Rogers LF. To see or not to see, that is the question: MR imaging of acute skeletal trauma. AJR Am J Roentgenol 2001;176(1):1.
2. Rogers LF. What is the role of MRI in acute skeletal trauma? AJR 2001;177:1245.
3. Maurer EJ, Kaplan PA, Dussault RG, Diduch DR, Schuett A, McCue FC, et al. Acutely injured knee: effect of MR imaging on diagnostic and therapeutic decisions. Radiology 1997;204(3):799-805.
4. Odgaard F, Tuxoe J, Joergensen U, Lange B, Lausten G, Brettlau T, et al. Clinical decision making in the acutely injured knee based on repeat clinical examination and MRI. Scand J Med Sci Sports 2002;12(3):154-62.
5. Gaebler C, Kukla C, Breitenseher M, Trattnig S, Mittlboeck M, Vecsei V. Magnetic resonance imaging of occult scaphoid fractures. J Trauma 1996;41(1):73-6.
6. Dorsay TA, Major NM, Helms CA. Cost-effectiveness of immediate MR imaging versus traditional follow-up for revealing radiographically occult scaphoid fractures. AJR Am J Roentgenol 2001;177(6):1257-63.
7. Breitenseher MJ, Metz VM, Gilula LA, Gaebler C, Kukla C, Fleischmann D, et al. Radiographically occult scaphoid fractures: value of MR imaging in detection [see comments]. Radiology 1997;203(1):245-50.
8. Carey J, Spence L, Blickman H, Eustace S. MRI of pediatric growth plate injury: correlation with plain film radiographs and clinical outcome. Skeletal Radiol 1998;27(5):250-5.
9. Lohman M, Kivisaari A, Kallio P, Puntila J, Vehmas T, Kivisaari L. Acute paediatric ankle trauma: MRI versus plain radiography. Skeletal Radiol 2001;30(9):504-11.
10. Rutt BK, Lee DH. The impact of field strength on image quality in MRI. J Magn Reson Imaging 1996;6(1):57-62.
11. Masciocchi C, Barile A, Satragno L. Musculoskeletal MRI: dedicated systems. Eur Radiol 2000;10(2):250-5.
12. Arbogast-Ravier S, Xu F, Choquet P, Brunot B, Constantinesco A. Dedicated low-field MRI: a promising low cost-technique. Med Biol Eng Comput 1995;33(5):735-9.
13. Barnett MJ. MR diagnosis of internal derangements of the knee: effect of field strength on efficacy. AJR Am J Roentgenol 1993;161(1):115-8.
14. Bradley WG. Future cost-effective MRI will be at high field. J Magn Reson Imaging 1996;6(1):63-6.
15. Chevrot A, Drapé J, Godefroy D, Dupont A. Dedicated MRI, emergency, cost-effectiveness. Abstract Industrial symposium 1997.
16. Davies A. Artoscan versus superconductive: the bottomline. Abstract Industrial symposium 1997.
17. Hayes E. Low-field and dedicated MRI gain credibility. Diagn Im Eur 1997(may):36-49.
18. Kersting-Sommerhoff B, Gerhardt P, Golder W, Hof N, Riel KA, Helmberger H, et al. [MRI of the knee joint: first results of a comparison of 0,2-T specialized system and 1,5-T high field strength magnet]. Rofo Fortschr Geb Rontgenstr Neuen Bildgeb Verfahr 1995;162(5):390-5.

19. Kersting-Sommerhoff B, Hof N, Lenz M, Gerhardt P. MRI of peripheral joints with a low-field dedicated system: a reliable and cost-effective alternative to high-field units? Eur Radiol 1996;6(4):561-5.
20. Mastantuono M, Passariello R. MRI evolution: from total body to "dedicated" systems. Abstract Industrial symposium 1997.
21. Parizel PM, Dijkstra HA, Geenen GP, Kint PA, Versteylen RJ, van Wiechen PJ, et al. Low-field versus high-field MR imaging of the knee: a comparison of signal behaviour and diagnostic performance [see comments]. Eur J Radiol 1995;19(2):132-8.
22. Passariello R. Niche MR unit shows promise in extremities. Diagn Imaging Europe 1995(march/apr):17-20.
23. Passariello R, Mastantuono M, Satragno L. Magnetic resonance imaging of the extremities using dedicated equipment: diagnostic and economic considerations. Acad Radiol 1996;3(Suppl 1):S141-4.
24. Peterfy CG, Roberts T, Genant HK. Dedicated extremity MR imaging. An emerging technology. Radiol Clin North Am 1997;35(1):1-20.
25. Rand T, Imhof H, Turetschek K, Schneider B, Vogele T, Gabler C, et al. Comparison of low field (0.2T) and high field (1.5T) MR imaging in the differentiation of torned from intact menisci. Eur J Radiol 1999;30(1):22-7.
26. Brouwer WB, Schut FT. Priority care for employees: a blessing in disguise? Health Econ 1999;8(1):65-73.
27. Rutten FFH, van Ineveld BM, van Ommen R, van Hout BA, Huijsman R. Kostenberekening bij Gezondheidszorgonderzoek. Utrecht: Jan van Arkel; 1993.
28. Arlen J. Tort damages. In: Bouckaert B, De Geest G, editors. Encyclopedia of Law and Economics. Ghent: University of Ghent; 1998.



MR Imaging of the Menisci and Cruciate Ligaments: A Systematic Review

Edwin H. G. Oei, MSc [1,2], Jeroen J. Nikken, MD [1,2],
Antonia C. M. Verstijnen, Msc [1,2], Abida Z. Ginai, MD, PhD [1],
M. G. Myriam Hunink, MD, PhD [1,2]

From the Program for the Assessment of Radiological Technology (ART Program).
Department of Radiology [1], Department of Epidemiology and Biostatistics [2],
Erasmus MC, University Medical Center,
Rotterdam. the Netherlands.

Radiology 2003; 226:837-848

Abstract

Purpose
To systematically review and synthesize published data on the diagnostic performance of magnetic resonance (MR) imaging of the menisci and cruciate ligaments and to assess the effect of study design characteristics and magnetic field strength on diagnostic performance.

Materials and Methods
Articles published between 1991 and 2000 were included if at least 30 patients were studied, arthroscopy was the reference standard, the magnetic field strength was reported, positivity criteria were defined, and the absolute numbers of true-positive, false-negative, true-negative, and false-positive results were available or derivable. Pooled weighted and summary receiver operating characteristic (ROC) analyses were performed for tears of both menisci and both cruciate ligaments separately and for the four lesions combined, by using random effects models. Differences were assessed according to lesion type.

Results
Twenty-nine of 120 retrieved articles were included. Pooled weighted sensitivity was higher for medial meniscal tears than that for lateral meniscal tears. However, pooled weighted specificity for the medial meniscus was lower than that for the lateral meniscus. In summary ROC analyses performed per lesion, various study design characteristics were found to influence diagnostic performance. Higher magnetic field strength significantly improved discriminatory power only for anterior cruciate ligament tears. When all lesions were combined in one overall summary ROC analysis, magnetic field strength was a significant but modest predictor of diagnostic performance.

Conclusion
Diagnostic performance of MR imaging of the knee is different according to lesion type and is influenced by various study design characteristics. Higher magnetic field strength modestly improves diagnostic performance, but a significant effect was demonstrated only for anterior cruciate ligament tears.

Introduction

Since its introduction for clinical use in the mid-1980s, the role of magnetic resonance (MR) imaging in the diagnosis of knee lesions has been established. MR imaging has proved reliable and safe and offers advantages over diagnostic arthroscopy, which is currently regarded as the reference standard for the diagnosis of internal derangements of the knee. Arthroscopy is an invasive procedure with certain risks and discomfort for the patient and is preferably performed only for treatment purposes, provided that alternative noninvasive diagnostic modalities such as MR imaging are available (1,2).

Results of numerous diagnostic studies have been published in which MR imaging and arthroscopy of the knee were compared, and most have shown good diagnostic performance in detecting lesions of the menisci and cruciate ligaments. As a result, MR imaging has been increasingly used in the diagnostic work-up of knee lesions, but the high costs of purchasing and maintaining a high-field-strength MR imager and the often limited availability of such a unit has restricted its widespread use for this purpose. The use of middle- and low-field-strength MR imagers has created the possibility of using MR imaging more routinely in the diagnostic work-up of knee disorders at a lower cost. In addition, low-field-strength dedicated extremity MR imagers have been designed specifically for extremity imaging, and because of their compact size and low field strength, costs can be kept relatively low compared with those of middle- and high-field-strength imagers. The reported diagnostic performance of low-field-strength MR imaging is variable, however, which raises the question of whether low-field-strength MR imaging can reliably replace middle- and high-field-strength MR imaging for evaluation of knee lesions.

With the vast number of articles on MR imaging of knee lesions with a wide range in study design, imaging techniques, and results, it is difficult to get a good idea of the diagnostic performance of MR imaging and the factors that influence its accuracy. We found some limited review articles in which investigators attempted to summarize published results of MR imaging of the knee (2,3). To our knowledge, however, no systematic review with a meta-analysis of the diagnostic performance of MR imaging of the knee has been published to date.

The purpose of this study was to systematically review and synthesize the published data on the diagnostic performance of MR imaging of the knee, focusing on tears of the menisci and cruciate ligaments, and to assess the effect of study design characteristics and magnetic field strength on diagnostic performance. We followed published guidelines for conducting diagnostic meta-analyses (4).

Materials and Methods

Selection of Articles

We conducted a MEDLINE search of the English-language literature to identify original articles published between January 1991 and December 2000 on the diagnostic performance of MR imaging of knee lesions. Combinations of the following search terms were used: "magnetic resonance imaging", "knee", "meniscus", "cruciate ligament", and "arthroscopy". All articles that could not be excluded definitively on the basis of the title and abstract were retrieved in full text. We used the criteria for inclusion and exclusion listed in Table 1. The bibliographies of the original articles were screened to obtain additional references.

TABLE 1. Criteria for Inclusion and Exclusion of Articles

Criteria type	Criteria
Inclusion	Articles in the English language
	MR imaging used to depict lesions of the medial or lateral meniscus, ACL, or PCL
	At least 30 patients examined
	Findings at arthroscopy used as reference standard
	Magnetic field strength reported
	Positivity criteria for MR imaging defined
	Absolute numbers of true-positive, false-negative, true-negative, and false-positive results available or derivable
Exclusion	Patient population explicitly stated to consist of infants or adolescents
	MR imaging used for postoperative evaluation
	Case-control study
	Results reported for only the medial and lateral meniscus combined
	Various magnetic field strengths used
	Only the diagnostic value of specific features and indirect signs of knee lesions at MR imaging, such as the empty notch sign, anterior tibial subluxation, or bone bruise, assessed

For the articles that were excluded because absolute numbers or positivity criteria were lacking, we tried to contact the corresponding author to request additional information. Authors were also contacted if there were indications that in more than one article, they reported on overlapping patient populations. If the authors did not respond, we chose the article that provided the most unequivocal and clear data.

Data Extraction
Two authors independently extracted results and characteristics of the study and MR imaging technique for each included article. For this purpose, a standardized data extraction sheet was used. Readers were not blinded to information about the authors, author affiliation, and journal name, since this has been shown to be unnecessary (5). The extracted study characteristics included publication year, country of origin, setting (academic or community hospital), patient characteristics, and aspects of study design, such as prospective versus retrospective design, inclusion of consecutive patients, and blinding. We also assessed the possibility of verification bias, which arises if patients are selected to undergo the reference test (arthroscopy) on the basis of the outcome of the test being evaluated (MR imaging), giving rise to an overestimation of sensitivity and an underestimation of specificity (6).
Three levels of likelihood were used to classify each study with respect to verification bias. We used "yes" if MR imaging was clearly used as a screening tool for arthroscopy. The category "possible" was assigned if, for unknown or unreported reasons, not all patients underwent both MR imaging and arthroscopy. We assumed that verification bias was absent if both MR imaging and arthroscopy were performed in all consecutive patients that were included in the study. Regarding the characteristics of MR imaging technique, we extracted the magnetic field strength of the MR imager and the number and type of MR imaging sequences. From the articles in which investigators tabulated the results for different readers, we extracted the data of the first reader, unless the level of experience was explicitly stated to be different among the readers. In the latter case, we extracted the results of the most experienced reader. If results were tabulated for multiple observations per reader, we extracted the data of the first observation. Some articles reported separately the results obtained with multiple MR imagers, which we treated as individual studies if the MR imagers had different magnetic field strengths. If results were reported for multiple MR imaging sequences or techniques with the same MR imager, we extracted the data of the technique that was recommended by the authors.
All discrepancies between our two data extractors were recorded. To determine the level of agreement, κ statistics were calculated for categorical variables, and Spearman correlation coefficients (r values) were calculated for continuous variables. To resolve discrepancies, a third data extractor assessed all discrepant items, and the majority opinion was used for analysis.

Data Analysis

Calculation of the natural logarithm of the diagnostic odds ratio. - For every lesion in each study, we calculated the natural logarithm of the diagnostic odds ratio *(OR) (D)*, which represents a summary measure of the diagnostic performance or discriminatory power. This value is the measure of interest in summary receiver operating characteristic (ROC) analyses and is calculated as follows:

$$D = \ln \{ [(TP + 0.5) (TN + 0.5)] / [(FP + 0.5) (FN + 0.5)] \},$$

where TP is the true-positive value, TN is the true-negative value, FP is the false-positive value, and FN is the false-negative value. Before calculating *D*, we first added 0.5 to each value to avoid undefined values of *D* and its variance resulting from zero values of each (7).

Assessment of publication bias. - We first evaluated the presence of publication bias, which could potentially arise if studies with positive results are more likely to be published than are those with negative results (8). Publication bias can be detected by constructing a funnel plot, in which the number of units measured (the number of knees) is plotted against the measure of interest (in our case, the natural logarithm of the diagnostic OR). In the absence of publication bias, the funnel plot shows a symmetric funnel-shaped distribution, whereas the distribution is asymmetric and skewed if publication bias is involved (9). Symmetry and shape of the funnel plots were judged by means of visual inspection.

Pooled weighted analyses. - To obtain a crude estimate of diagnostic performance of MR imaging in depicting tears of the medial and lateral menisci, anterior cruciate ligaments (ACLs), and posterior cruciate ligaments (PCLs), we performed a pooled weighted analysis per lesion, weighting with the reciprocal of the variance of each study. First, we tested for heterogeneity in effect size among studies, the result of which determines if a fixed or random effects model should be used (10,11). Pooled weighted sensitivity, specificity, and natural logarithm of the diagnostic OR (and 95% CIs) were then calculated by using a random effects model. We regarded two estimates as significantly different if their 95% CIs did not overlap.

Next, pooled weighted analyses for the different lesions were performed for various categories of magnetic field strength separately to obtain an idea of the influence of magnetic field strength on diagnostic performance. The range of field strengths was divided into either three categories (higher than 1.0 T, 0.5 - 1.0 T, or lower than 0.5 T) or two categories (1.0 T and higher vs lower than 1.0 T).

TABLE 2. Characteristics of Included Studies

Study	Public. Year	Country of Origin	Hospital Type	Study Type	Blinding	Verification Bias	Consecutive Patients	Mean Age (y)
Boeree et al (25)	1991	UK	Community	Retrospective	NR	Yes	NR	34
Fischer et al (26)	1991	USA	NR	NR	NR	Yes	NR	37
Kelly et al (27)	1991	USA	Community	Retrospective	NR	Possible	NR	37
Niitsu et al (28)	1991	Japan	Academic	NR	NR	No	NR	25
Araki et al (29)	1992	Japan	NR	NR	NR	Yes	Yes	31
Gluckert et al (30)	1992	Germany	Academic	Prospective	NR	No	NR	33
Grevitt et al (31)	1992	UK	Community	Prospective	Yes	No	Yes	36
Gross et al (32)	1992	USA	Academic	Retrospective	NR	Yes	Yes	NR
Heron and Calvert (33)	1992	UK	Community	Prospective	No*	Yes	No	38
Barnett (34)	1993	USA	Community	Retrospective	NR	Yes	Yes	NR
De Smet et al (35)	1993	USA	Academic	Retrospective	NR	Yes	Yes	29
Chan et al (36)	1994	USA	Academic	Retrospective	NR	Possible	Yes	37
Gentili et al (37)	1994	USA	NR	Retrospective	NR	Possible	Yes	31
Kinnunen et al (38)	1994	Finland	Academic	Prospective	Yes	No	Yes	36
LaPrade et al (39)	1994	USA	Academic	Prospective	NR	No	NR	34
Chen et al (40)	1995	Taiwan	Academic	Retrospective	NR	Possible	NR	NR
Justice and Quinn (41)	1995	USA	NR	Prospective	No	Possible	Yes	NR
Lerman et al (42)	1995	USA	NR	Retrospective	NR	Yes	NR	37
Barry et al (43)	1996	USA	Academic	Retrospective	NR	Possible	NR	29.7
Lundberg et al (44)	1996	Sweden	Academic	Prospective	No	No	Yes	26
Bui-Mansfield et al (46)	1997	USA	Community	Prospective	No	No	Yes	31
Cheung et al (47)	1997	USA	Academic	Retrospective	NR	Possible	No	38
Franklin et al (48)	1997	USA	NR	Prospective	Yes	No	Yes	40
Rappeport et al (49)	1997	Denmark	Academic	Prospective	Yes	No	NR	NR
Weinstabl et al (50)	1997	Austria	Academic	Prospective	NR	Yes	Yes	34.9
Ha et al (51)	1998	USA	Academic	Retrospective	NR	Possible	NR	40
Riel et al (45)	1999	Germany	Academic	Prospective	NR	No	NR	36
Cotten et al (52)	2000	France	Community	Prospective	NR	Possible	NR	34
Elvenes et al (53)	2000	Norway	Academic	NR	NR	Possible	NR	32

Note.- Numbers in parentheses are reference numbers. NR = not reported, * No in 85 patients, Yes in 15 patients.

Summary ROC Analysis per Lesion Type

We subsequently performed summary ROC analysis for each of the lesions separately, which has advantages over pooled weighted analysis. First, the true- and false-positive rates for the different diagnostic studies can be summarized and synthesized, with adjustment for different positivity criteria among studies (12-14). Different positivity criteria exist among studies when institutions use different thresholds for labeling a test result as positive. For example, in one study, an area of high signal intensity in the course of the cruciate ligament may be considered indicative of a rupture, whereas in another study, investigators may require the absence of intact cruciate ligament margins as an additional finding before labeling the test result positive. Second, by using summary ROC analysis, one can identify and adjust for variables that have an influence on diagnostic performance. We applied a random effects model, which accounted for the residual interstudy heterogeneity, which may be present even after adjustment for characteristics such as population size, patient age and sex, positivity criteria, type of MR imager, or verification bias. A random effects regression model was developed in which the natural logarithm of the diagnostic OR of each study was the dependent variable. As the independent variable, a measure of the positivity criteria (S) of the study was calculated as follows:

$$S = \ln \{[(TP + 0.5)(FP + 0.5)] / [(TN + 0.5)(FN + 0.5)]\},$$

where TP is the true-positive value, FP is the false-positive value, TN is the true-negative value, and FN is the false-negative value. We added 0.5 to each value as before to prevent undefined values of the positivity criteria. By using this method, one adjusts for the variations in positivity criteria that are used implicitly or explicitly in different studies and that are assumed to influence diagnostic performance.

Other variables that influence diagnostic performance were identified by adding them to the regression model and by assessing their regression coefficient and influence on the model. The variable, depending on the type, can be included either directly as a continuous variable or as a dummy variable. In this bivariate summary ROC analysis, we considered variables as explanatory if they were statistically significant in the regression model ($P < 0.05$) or if the regression coefficient was at least 1.0 for dummy variables or 1.0 over the range of the variable values. Variables with a P value between 0.05 and 0.10 were retained in the model if their inclusion decreased the method-of-moments τ^2 estimate by at least 10% compared with that in the univariate model. The method-of-moments τ^2 calculation provides a measure of between-study variance and is higher if there is more heterogeneity among studies and zero if studies are homogeneous.

TABLE 3. MR Imaging Technique and Characteristics of Included Studies: Medial Meniscal Tears

Study	Publication Year	Magnetic Field Strength (T)	No. of Sequences	No. of Patients	TP	FP	TN	FN
Boeree et al (25)	1991	0.5	2	129	58	6	63	2
Fischer et al (26)	1991	1.5	3	483	270	33	170	10
Fischer et al (26)	1991	0.35	3	177	70	11	83	13
Fischer et al (26)	1991	1.5	3	86	37	2	45	2
Fischer et al (26)	1991	0.6	3	373	32	9	27	5
Fischer et al (26)	1991	0.35	3	36	8	11	15	2
Kelly et al (27)	1991	0.5	3	60	33	6	20	1
Araki et al (29)	1992	1.5	1	40	19	2	18	1
Gluckert et al (30)	1992	1.5	1	80	34	3	42	1
Grevitt et al (31)	1992	0.2	2	55	23	3	27	2
Heron and Calvert (33)	1992	1.5	3	100	44	3	52	1
Barnett (34)	1993	0.5	6	118	71	4	38	5
De Smet et al (35)	1993	1.5	2	200	100	13	79	8
Kinnunen et al (38)	1994	0.1	3	33	7	5	20	1
LaPrade et al (39)	1994	1.0	5	72	34	1	37	0
Chen et al (40)	1995	1.5	6	50	13	5	30	2
Justice and Quinn (41)	1995	1.5	4	561	360	16	170	15
Lundberg et al (44)	1996	1.5	3	69	14	17	33	5
Bui-Mansfield et al (46)	1997	1.5	4	50	18	1	29	2
Cheung et al (47)	1997	1.5	5	289	127	23	123	16
Franklin et al (48)	1997	0.2	6	35	19	0	13	3
Rappeport et al (49)	1997	0.1	3	47	12	9	24	2
Weinstabl et al (50)	1997	1.5	2	75	48	4	22	1
Riel et al (45)	1999	0.2	6	244	106	3	127	8
Cotten et al (52)	2000	0.2	4	90	50	0	33	7
Cotten et al (52)	2000	1.5	4	90	51	0	33	6
Elvenes et al (53)	2000	0.5	2	41	15	6	20	0

Note. - Numbers in parentheses are reference numbers. Data are number of patients, unless otherwise indicated. TP = true-positive value, FP = false-positive value, TN = true-negative value, FN = false-negative value.

We assessed as potential predictors of diagnostic performance the effect of publication year (continuous variable), country of origin (in North America vs other), type of hospital (academic, community, or not reported), mean age (continuous variable, 35 years or older vs younger than 35 years), inclusion of consecutive patients (yes, no, or not reported), prospective versus retrospective study

design, verification bias (yes, no, or possible), blinding of the arthroscopist to the MR imaging result (yes, no, or not reported), magnetic field strength of the MR imager (continuous variable, 1.5 T vs lower than 1.5 T), and number of MR sequences used (continuous variable). In five articles, mean age was not reported, and in the analyses in which mean age was examined, regression analysis was performed for the subset of studies with age available. In all studies, the radiologist was blinded to the arthroscopic findings because MR imaging was always performed prior to arthroscopy. Therefore, only the effect of blinding of the arthroscopist was assessed in this study.
Subsequently, we performed multivariate summary ROC analysis, in which the explanatory variables previously identified in the bivariate analysis were included one by one in a stepwise forward-selection regression model. We started with the variable that decreased the method-of-moments τ^2 estimate the most and kept it in the model on the basis of the same criteria as those used in the bivariate summary ROC analysis. We always retained the measure of positivity criteria (S) in the model, since a difference in positivity criteria among studies is a key concept of summary ROC analysis. Even though the positivity criteria as defined in the articles seemed largely the same, interpretation of MR images of the knee may vary in subtle ways in different hospitals and among different radiologists. This yielded the final model for each lesion from which multivariate summary ROC curves were plotted, adjusted for significant covariables that were set to the mean values or values indicating the ideal study design, as appropriate.

Overall Summary ROC Analysis of All Lesions

For the purpose of increasing statistical power and precision and comparing the results with those of the separate analyses per lesion, we also analyzed all four lesions in a single model. To compare diagnostic performance among lesion types, we created dummy variables to code if the lesion of interest was a meniscal versus cruciate ligament tear.
The effect of each significant predictor in the final models is reflected by its regression coefficient. A positive regression coefficient indicates better discriminatory power of MR imaging in studies with that predictor, compared with that in studies without the corresponding characteristic. A negative regression coefficient indicates reduced diagnostic performance in studies with that characteristic. Finally, we calculated relative diagnostic ORs of all predictors in the multivariate models by taking the antilogarithm of the regression coefficient. A relative diagnostic OR can be interpreted as the diagnostic performance of a test in studies with a certain characteristic, relative to its performance in studies without that corresponding feature.

TABLE 4. MR Imaging Technique and Characteristics of Included Studies: Lateral Meniscal Tears

Study	Publication Year	Magnetic Field Strength (T)	No. of Sequences	No. of Patients	TP	FP	TN	FN
Boeree et al (25)	1991	0.5	2	127	25	2	99	1
Fischer et al (26)	1991	1.5	3	513	81	20	378	34
Fischer et al (26)	1991	0.35	3	192	20	5	148	19
Fischer et al (26)	1991	1.5	3	89	18	3	66	2
Fischer et al (26)	1991	0.6	3	77	9	8	55	5
Fischer et al (26)	1991	0.35	3	36	6	2	27	1
Kelly et al (27)	1991	0.5	3	60	19	5	34	2
Araki et al (29)	1992	1.5	1	40	13	0	27	0
Gluckert et al (30)	1992	1.5	1	80	12	0	68	0
Grevitt et al (31)	1992	0.2	2	55	8	1	45	1
Heron and Calvert (33)	1992	1.5	3	100	17	5	77	1
Barnett (34)	1993	0.5	6	118	21	3	89	5
De Smet et al (35)	1993	1.5	2	200	48	12	130	10
Kinnunen et al (38)	1994	0.1	3	33	1	1	28	3
LaPrade et al (39)	1994	1.0	5	72	14	1	51	6
Chen et al (40)	1995	1.5	6	50	17	3	27	3
Justice and Quinn (41)	1995	1.5	4	561	134	8	390	29
Lundberg et al (44)	1996	1.5	3	69	13	7	36	13
Bui-Mansfield et al (46)	1997	1.5	4	50	9	0	35	6
Cheung et al (47)	1997	1.5	5	289	69	13	180	27
Franklin et al (48)	1997	0.2	6	35	8	0	26	1
Rappeport et al (49)	1997	0.1	3	47	2	1	41	3
Weinstabl et al (50)	1997	1.5	2	75	17	1	56	1
Riel et al (45)	1999	0.2	6	244	38	7	191	8
Cotten et al (52)	2000	0.2	4	90	24	6	55	5
Cotten et al (52)	2000	1.5	4	90	25	3	58	4
Elvenes et al (53)	2000	0.5	2	41	2	4	32	3

Note. - Numbers in parentheses are reference numbers. Data are number of patients, unless otherwise indicated. TP = true-positive value, FP = false-positive value, TN = true-negative value, FN = false-negative value.

Thus, a relative diagnostic OR greater than 1.0 indicates that studies with that characteristic yield better diagnostic performance of MR imaging than that instudies without the corresponding feature, whereas a relative diagnostic OR less than 1.0 indicates reduced discriminatory power in studies with that characteristic.

All analyses were performed by using STATA (versions 6.0 and 7.0; Stata, College Station, Tex) and SPSS for Windows (version 9.0.0; SPSS, Chicago, Ill) software.

Sensitivity Analyses

To assess the dependence of the five final multivariate models on the results of individual studies, we performed the jackknife type of sensitivity analyses for each model. By using this method, one can determine the contribution of the individual studies to the overall results by performing multiple summary ROC analyses with each article excluded in turn (the jackknife method).

RESULTS

Literature Search and Data Extraction

Our MEDLINE search resulted in 804 articles, of which 120 were retrieved after we evaluated the titles and abstracts. Eighty nine articles were excluded because *(a)* the article was a review or descriptive article without original data on diagnostic performance (*n* = 20), *(b)* fewer than 30 patients were studied (*n* = 9), *(c)* magnetic field strength was not reported or was variable (*n* = 13), *(d)* absolute numbers were not available and could not be derived (*n* = 19), *(e)* no positivity criteria at MR imaging were reported (*n* = 3), *(f)* results were reported for the medial and lateral meniscus combined (*n* = 6), *(g)* only the value of specific indirect signs or features of knee lesions at MR imaging was assessed (*n* = 15), or *(h)* the patient population was suspected to overlap with that of another study (*n* = 4). Furthermore, we excluded one article in which patients and random control subjects were retrospectively selected on the basis of the presence or absence of a definite complete ACL tear (case-control design) without accounting for all patients in the study period (15). Another article was excluded in which the whole knee was used as the unit of analysis (16). All excluded articles, together with the reasons for exclusion, are listed in the table that is available as supplemental material on the *Radiology* website:

Table E1, *radiology .rsnajnls.org/cgi/content/full/2263011892 /DC1.*

In this table, each article is classified according to the reason why the article was excluded from our analysis, although multiple reasons may be applicable to one article. We attempted to contact 16 authors for more information, but this did not result in additional articles for inclusion. Thus, 29 articles were included in this meta-analysis, which comprised 27 studies on both menisci, 23 studies on ACL tears, and 12 studies on PCL tears. The articles, together with the corresponding study characteristics, are listed in Table 2. Results for the menisci and cruciate ligaments per study, as well as features of the MRimaging technique, are listed in Tables 3-5.

TABLE 5. MR Imaging Technique and Characteristics of Included Studies: Cruciate Ligament Tears

Study	Publication Year	Magnetic Field Strength (T)	No. of Sequences	No. of Patients	TP	FP	TN	FN
ACL complete tear								
Boeree et al (25)	1991	0.5	2	133	32	11	89	1
Fischer et al (26)	1991	1.5	3	530	100	30	396	4
Fischer et al (26)	1991	0.35	3	193	28	8	157	0
Fischer et al (26)	1991	1.5	3	90	17	3	70	0
Fischer et al (26)	1991	0.6	3	79	13	5	56	5
Fischer et al (26)	1991	0.35	3	41	5	7	27	2
Kelly et al (27)	1991	0.5	3	60	7	3	49	1
Niitsu et al (28)	1991	1.5	1	52	20	3	21	8
Gluckert et al (30)	1992	1.5	1	80	18	1	60	1
Grevitt et al (31)	1992	0.2	2	55	8	4	43	0
Heron and Calvert (33)	1992	1.5	3	99	20	4	73	2
Barnett (34)	1993	0.5	6	118	26	3	89	0
Chan et al (36)	1994	1.5	4	120	19	6	93	2
Gentili et al (37)	1994	1.5	4	89	48	1	39	1
Kinnunen et al (38)	1994	0.1	3	33	5	4	23	1
Chen et al (40)	1995	1.5	6	50	22	3	24	1
Lerman et al (42)	1995	1.5	1	47	23	0	22	2
Barry et al (43)	1996	0.3	4	63	25	4	31	3
Bui-Mansfield et al (46)	1997	1.5	4	50	20	3	27	0
Rappeport et al (49)	1997	0.1	3	47	6	3	37	1
Ha et al (51)	1998	1.5	5	217	54	4	157	2
Cotten et al (52)	2000	0.2	4	90	14	2	73	1
Cotten et al (52)	2000	1.5	4	90	14	0	75	1
PCL complete tear								
Fischer et al (26)	1991	1.5	3	537	2	3	530	2
Fischer et al (26)	1991	0.35	3	198	3	0	195	0
Fischer et al (26)	1991	1.5	3	93	2	1	90	0
Fischer et al (26)	1991	0.6	3	80	1	0	79	0
Fischer et al (26)	1991	0.35	3	41	0	1	40	0
Niitsu et al (28)	1991	1.5	1	52	4	1	47	0
Gluckert et al (30)	1992	1.5	1	80	0	0	80	0
Gross et al (32)	1992	0.3	1	203	13	0	190	0
Heron and Calvert (33)	1992	1.5	3	99	1	0	98	0
Kinnunen et al (38)	1994	0.1	3	33	0	1	32	0
Chen et al (40)	1995	1.5	6	50	4	0	46	0
Bui-Mansfield et al (46)	1997	1.5	4	50	1	8	41	0

Note. - Numbers in parentheses are reference numbers. Data are number of patients, unless otherwise indicated. TP = true-positive value, FP = false-positive value, TN = true-negative value, FN = false-negative value.

We extracted 379 items from the articles; 46 discrepancies occurred between our two data extractors (Table 6), which were resolved by the third extractor.

TABLE 6. Analyzed Variables and Measures of Agreement between Two Data Extractors for 379 Items in 29 Articles

Variable	No. of Discrepancies*	Spearman *r* Value	κValue
Publication year	0 (0)	1.00	NA
Country of origin	0 (0)	NA	1.00
Academic or community hospital	5 (17)	NA	0.70
Mean age	0 (0)	1.00	NA
Consecutive patients	8 (28)	NA	0.54
Prospective or retrospective study design	6 (21)	NA	0.67
Verification bias	8 (28)	NA	0.59
Blinding of arthroscopist to MR imaging findings	4 (14)	NA	0.71
Magnetic field strength	0 (0)	1.00	NA
Number of pulse sequences	8 (28)	0.80	NA
Natural logarithm of diagnostic OR	7 (8)	0.96	NA
Overall	46 (12)	NA	NA

Note. - The Spearman *r* value was calculated for continuous variables, the k value for categoric variables. NA = not applicable.* Numbers in parentheses are percentages.

Depending on the variable, agreement between the two data extractors ranged from moderate to almost perfect (*r* value range, 0.80-1.00; κ value range, 0.54-1.00).

Publication Bias

At visual inspection, the funnel plots for the menisci showed an almost perfectly funnel-shaped distribution (Fig 1), suggesting that publication bias is unlikely. For both cruciate ligaments, the funnel plots were almost funnel shaped, although not perfectly so (Fig 2), but there were sufficient small studies with a low diagnostic performance to suggest the absence of substantial publication bias.

Pooled Weighted Analysis

The test for heterogeneity showed significant results for all lesions, with the exception of PCL tears. We therefore used a random effects model, which can accommodate both heterogeneous and homogeneous effect sizes among studies (10,11).

TABLE 7. Pooled Weighted Results per Type of Lesion (random effects model)

Lesion	Pooled Weighted Sensitivity (%)	Pooled Weighted Specificity (%)	Pooled Weighted D^* Value
Medial meniscal tear	93.3 (91.7, 95.0)	88.4 (85.4, 91.4)	4.38 (3.86, 4.89)
Lateral meniscal tear	79.3 (74.3, 84.2)	95.7 (94.6, 96.8)	4.07 (3.60, 4.54)
ACL complete tear	94.4 (92.3, 96.6)	94.3 (92.7, 95.9)	4.90 (4.28, 5.53)
PCL complete tear	91.0 (83.2, 98.7)	99.4 (98.9, 99.9)	5.42 (4.40, 6.45)

Note. - Numbers in parentheses are 95% CIs. * D = Natural logarithm of the diagnostic OR.

For the diagnosis of meniscal tears (Table 7), pooled weighted sensitivity was higher for the medial meniscus (93.3% [95% CI: 91.7, 95.0]) than that for the lateral meniscus (79.3% [95% CI: 74.3, 84.2]), whereas pooled weighted specificity for medial meniscal tears (88.4% [95% CI: 85.4, 91.4]) was lower than that for lateral meniscal tears (95.7% [95% CI: 94.6, 96.8]). There was no statistically significant difference in sensitivity for complete ACL tears versus PCL tears (Table 7) (94.4% [95% CI: 92.3, 96.6] and 91.0% [95% CI: 83.2, 98.7], respectively), but the specificity for ACL tears was lower than that for PCL tears (94.3% [95% CI: 92.7, 95.9] and 99.4% [95% CI: 98.9, 99.9], respectively). No significant difference was demonstrated in the pooled natural logarithms of the diagnostic ORs for the different lesions. The results of the separate pooled weighted analyses for various categories of magnetic field strengths (not tabulated) suggested a modest trend toward better diagnostic performance for higher-field-strength categories. None of the differences were found to approach statistical significance, however, and the CIs were all extremely wide because of a limited number of studies per category of magnetic field strength.

Summary ROC Analysis per Type of Lesion

In the multivariate summary ROC analysis for the medial meniscus, blinding of the arthroscopist was the only predictor of diagnostic performance. For the lateral meniscus, mean age was the only significant variable in the final model (Table 8). Publication year, mean age, and magnetic field strength were predictors in the multivariate model for ACL tears. The final model for PCL tears consisted of publication year, academic hospital setting, verification bias, and number of MR sequences. The summary ROC curves based on the final regression models, together with the results of the individual studies, are plotted in Figure 3.

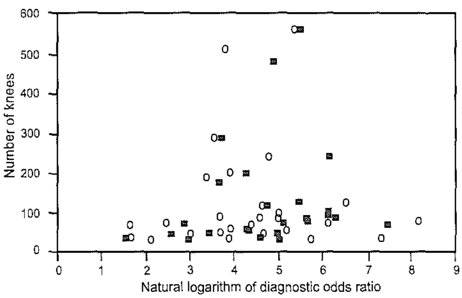


Figure 1. Funnel plot for the menisci in which the number of knees is plotted against the discriminatory power of MR imaging (natural logarithm of the diagnostic OR). For both medial and lateral menisci, the distribution of data points appears to be fairly funnel shaped and symmetric, indicating that publication bias is unlikely. Filled box = results for medial meniscus from individual studies, open point = results for lateral meniscus from individual studies.

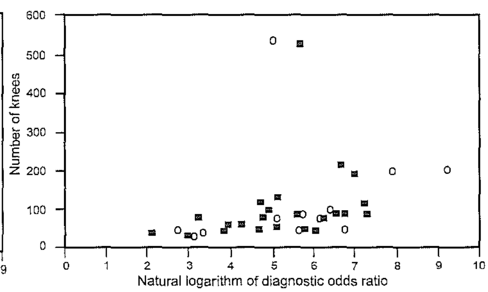


Figure 2. Funnel plot for the cruciate ligaments in which the number of knees is plotted against the discriminatory power of MR imaging (natural logarithm of the diagnostic OR). For both ACL and PCL lesions, the data points show a slightly skewed distribution. There is a considerable number of small studies with low diagnostic performance, however, suggesting the absence of substantial publication bias. Filled box = ACL results from individual studies, open point = PCL results from individual studies.

For the medial meniscus, we set the dummy variable for blinding to zero according to clinical practice, where arthroscopy is ideally performed on the basis of and with knowledge of the MR imaging findings. The summary ROC curve for lateral meniscal tears was plotted for a patient aged 30 years. For the ACL, the publication year was set to 1995, and the curve was plotted for a patient aged 30 years with use of a 1.0-T MR imager. The curve for PCL lesions was adjusted to publication in 1995, an academic hospital setting, use of three MR sequences, and the absence of verification bias. The adjusted summary ROC curves have different positions in ROC space (Fig 3), indicating that there are differences in diagnostic performance among the types of lesions. Whereas overall discriminatory power is not much different for the menisci, diagnostic performance is clearly better for the ACL compared with that for the PCL, since the curve for the ACL is located more toward the upper left corner.

Overall Summary ROC Analysis of All Lesions

In the multivariate summary ROC analysis with all lesions combined, the meniscal tear versus cruciate ligament dummy variable (relative diagnostic OR, 0.52 [95% CI: 0.26, 1.01]), mean age (relative diagnostic OR, 1.13 [95% CI: 1.03, 1.23]), and magnetic field strength (relative diagnostic OR, 1.97 [95% CI: 1.13, 3.42]) were statistically significant predictors of diagnostic performance (Table 8).

TABLE 8. Multivariate Summary ROC Models per Type of Lesion and for All Lesions Combined (random effects model)

Covariable	Regression Coefficient	Relative Diagnostic OR	*P* Value	Method-of-Moments τ^2
Medial meniscal tear (27 studies)				
Positivity criteria	−0.32 (−0.87, 0.23)	0.73 (0.42, 1.26)	0.25	1.22
Blinding arthroscopist to MR imaging findings	−1.09 (−2.68, 0.50)	0.34 (0.07, 1.65)	0.18	NA
Lateral meniscal tear (23 studies) *				
Positivity criteria	0.39 (−0.12, 0.90)	1.48 (0.89, 2.46)	0.14	0.69
Mean age	0.11 (−0.04, 0.25) †	1.12 (0.96, 1.28)	0.14	NA
ACL complete tear (20 studies) ‡				
Positivity criteria	0.64 (0.10, 1.18)	1.90 (1.11, 3.25)	0.02	0.18
Publication year	0.25 (0.08, 0.41) †	1.28 (1.08, 1.51)	0.004	NA
Mean age	0.12 (0.00, 0.24) †	1.13 (1.00, 1.27)	0.05	NA
Magnetic field strength	1.21 (0.40, 2.02) §	3.35 (1.49, 7.54)	0.003	NA
PCL complete tear (12 studies)				
Positivity criteria	0.22 (−0.72, 1.16)	1.25 (0.49, 3.19)	0.64	0.00
Publication year	0.21 (−0.92, 1.34) †	1.23 (0.40, 3.82)	0.72	NA
Academic hospital	1.89 (−0.68, 4.46)	6.62 (0.51, 86)	0.15	NA
Verification bias	4.23 (0.28, 8.18)	68.72 (1.32, 3,569)	0.04	NA
Number of MR sequences	−0.75 (−2.28, 0.78) ¶	0.47 (0.10, 2.18)	0.34	NA
All lesions combined (76 studies) #				
Positivity criteria	0.06 (0.15, 0.28)	1.06 (0.86, 1.32)	0.56	0.87
Meniscus vs cruciate ligament	0.66 (1.33, 0.01)	0.52 (0.26, 1.01)	0.05	NA
Mean age	0.12 (0.03, 0.21) †	1.13 (1.03, 1.23)	0.009	NA
Magnetic field strength	0.68 (0.12, 1.2) §	1.97 (1.13, 3.42)	0.02	NA

Note. – Numbers in parentheses are 95% CIs. NA = not applicable. A positive regression coefficient indicates better discriminatory power of MR imaging in studies with that characteristic compared with that in studies without the corresponding characteristic, and a negative regression coefficient indicates reduced diagnostic performance in studies with that characteristic. A relative diagnostic OR greater than 1.0 indicates better diagnostic performance of MR imaging in studies with that characteristic compared with that in studies without the corresponding feature, and a relative diagnostic OR less than 1.0 indicates decreased diagnostic performance in studies with that characteristic.

* Subset analysis including 23 of 27 studies with age reported.

† Value is the change per year.

‡ Value is the change per unit Tesla.

§ Subset analysis including 20 of 23 studies with age reported.

¶ Value is the change per additional sequence.

Subset analysis including 76 of 89 studies with age reported.

Although the dummy variable for meniscal versus cruciate ligament tears was of borderline significance, we kept it in the model because the nature of these lesions is totally different, and furthermore, it allowed comparison of the lesions.

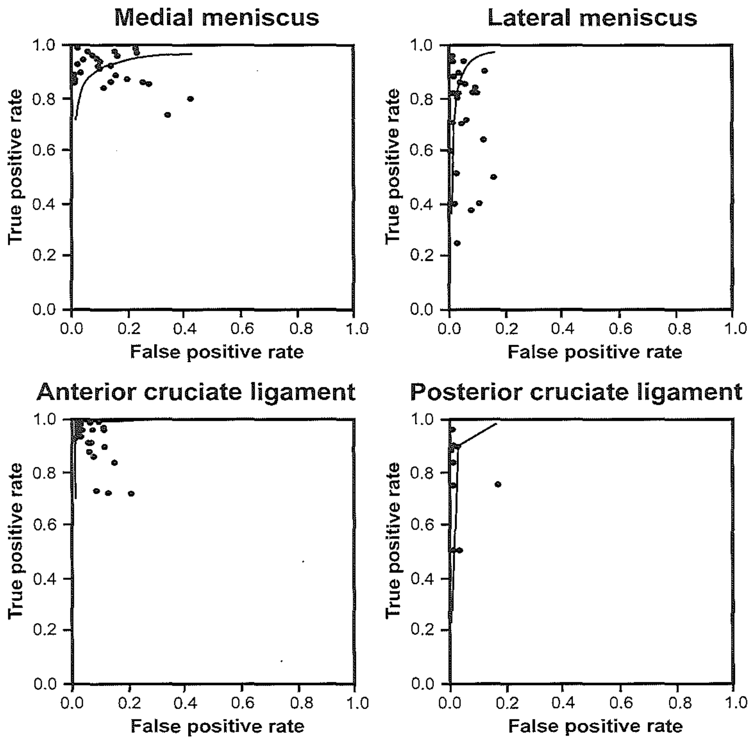


Figure 3. Summary ROC curves for the four types of lesions separately and the results of the individual studies (filled points) on the basis of the final models per lesion. Upper left: Summary ROC curve for medial meniscal tears. The dummy variable for blinding was set to zero, according to clinical practice, in which MR imaging is ideally performed before arthroscopy. Upper right: Summary ROC curve for lateral meniscal tears, adjusted to a patient aged 30 years. Lower left: Summary ROC curve for ACL tears, adjusted to publication in 1995, a patient aged 30 years, and a 1.0-T MR imager. Lower right: Summary ROC curve for PCL tears, adjusted for publication in 1995, academic hospital setting, three MR sequences in the imaging protocol, and absence of verifica-tion bias. The adjusted summary ROC curves demonstrate the differences in diagnostic performance among the lesion types. Whereas overall discriminatory power is not much different for the menisci, diagnostic performance is clearly better for the ACL compared with that for the PCL

In Figure 4, summary ROC curves are shown for menisci and cruciate ligaments separately. The curves were plotted with the mean age covariable set at 30 years for both 0.2- and 1.5-T MR imagers. Compared with that for the menisci, the curve for cruciate ligaments is further up toward the upper left corner, suggesting better discriminatory power. Similarly, the curves for high-field-strength MR imaging show better diagnostic performance compared with those for low-field-strength MR imaging.

Sensitivity Analyses

The jackknife sensitivity analyses, in which articles were excluded one by one from the final models, did not demonstrate any disproportionate influences of individual studies. In the models for ACL tears and the overall model with all lesions combined, the effect of magnetic field strength remained relatively stable in the sensitivity analysis (relative diagnostic OR ranges, 2.63-4.41 and 1.80-2.16, respectively), and this effect was always statistically significant.

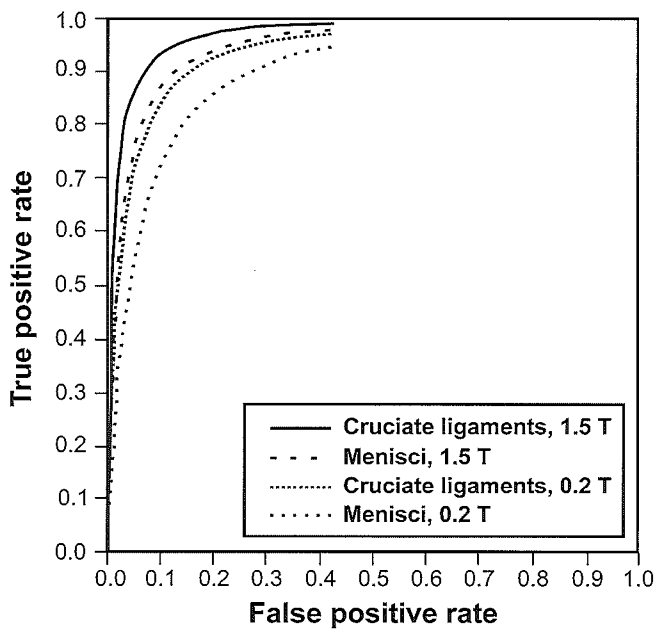


Figure 4. Multivariate summary ROC curves based on the final regression model with all lesions combined. Curves are shown for both menisci and both cruciate ligaments combined. Curves were adjusted to patients aged 30 years and both 0.2- and 1.5-T MR imagers. The curves for cruciate ligaments are further up toward the upper left corner, as are the curves for high-field-strength MR imaging, suggesting better discriminatory power compared with that of the menisci and low-field-strength MR imaging.

Discussion

In this systematic review and meta-analysis, attention was focused on tears of the menisci and cruciate ligaments, which are primarily caused by traumatic mechanisms. Our results confirm those of previous studies, which show MR imaging to be a highly accurate diagnostic tool for detecting tears of the menisci and cruciate ligaments. The results show that diagnostic performance is better for cruciate ligament tears than that for meniscal tears. With regard to the menisci, our results demonstrate that the sensitivity and specificity differ significantly for the medial and lateral meniscus.

Whereas MR imaging is more sensitive in the diagnosis of medial meniscal tears, the specificity is higher for lateral meniscal tears. However, the natural logarithm of the diagnostic OR, which is an overall measure of diagnostic performance that incorporates both sensitivity and specificity, is not significantly different for the two lesions, indicating that radiologists probably use different points along the same underlying ROC curve when evaluating the two lesions. The results confirm that findings of the two menisci are preferably considered separately in studies in which the diagnostic performance of MR imaging was assessed to avoid performance underestimation. Thus, our criterion to include only articles in which results for medial and lateral menisci were reported separately seems to be justified. Moreover, in clinical practice, it is usually a lesion of either the medial or lateral meniscus that is suspected, making diagnostic performance statistics for both menisci combined less meaningful. To increase statistical precision, a pooled weighted analysis with both menisci combined into one model is possible, provided that an appropriate technique (eg, summary ROC analysis) for combining lesions with different points on the same ROC curve is used, which is what we did. We acknowledge, however, that this type of pooled combined analysis may result in data for some patients being included more than once in the same regression model, which induced some dependence among observations.

With regard to the cruciate ligaments, we considered complete tears only, because these injuries are by far more serious than are partial ruptures. Whereas complete ruptures mostly necessitate intensive physical therapy or reconstructive surgery, partial tears usually do not require specific treatment. Apart from the menisci and cruciate ligaments, there are more derangements involving the knee that might be visualized with MR imaging. Articular cartilage lesions may also have a traumatic origin and cause serious symptoms. Except for full-thickness lesions (Ficat grade IV or V), however, the role of MR imaging in detecting articular cartilage defects has not been well established. Especially for lesions that are limited to half of the cartilage thickness (Ficat grade I-III), the diagnostic accuracy of MR imaging seems poor (17-19). We performed a MEDLINE search of articles on MR imaging of knee cartilage, but among the articles that we found, we considered the patient population, grading system, definition of disease, and regions studied· as too heterogeneous to justify a meta-analysis. For example, patient populations varied from patients who experienced trauma to patients with known osteoarthritis. Grading systems had a range of three to six categories, and positivity criteria varied from anything abnormal to full-thickness defects. Finally, in some studies, the knee cartilage was subdivided into regions, such as the lateral and medial femoral condyle or even the different facets of the patella, whereas in other articles, investigators considered the whole knee to be a unit of analysis.

To avoid missing important articles, we applied broad criteria for our MEDLINE search. This resulted in 804 retrieved references, most of which were clearly not suitable for inclusion in our systematic review and were excluded on the basis of either the title or the abstract. As an illustration, a substantial number of articles about other pathologic conditions in the knee or about postoperative MR imaging, cadaveric examinations, case reports, and even the meniscus in the temporomandibular joint were found among these 804 studies. After this initial selection, 120 articles were considered potentially eligible for inclusion, 20 of which were subsequently excluded because they did not present data on diagnostic performance. We applied criteria for inclusion and exclusion that are commonly used in evidence-based medicine, as well as criteria that were related to the aims of the present study and the methods that we intended to use. Studies with a sample size of fewer than 30 patients were excluded because small samples contribute little to the results of a meta-analysis. In addition, our purpose to assess the effect of magnetic field strength on diagnostic performance required that the magnetic field strength of the MR imager be reported in all included studies. Similarly, a summary ROC analysis is only possible if the absolute numbers of true-positive, false-negative, true-negative, and false-positive results are available or derivable.

Our literature search was limited to articles published between 1991 and 2000. We acknowledge that there have been relevant publications between 1985 and 1990, some of which would have fulfilled the inclusion criteria of this meta-analysis. However, the question arises whether the MR imaging techniques used in the earlier period can be compared with the improved techniques and sequences of more recent years. Furthermore, we limited our literature search to articles published in the English language. It has been shown that inclusion of only English-language articles does not influence the results of a meta-analysis (20).

Moreover, the decision as to which other languages to include will, in our view, always be based on highly arbitrary and geographically dependent criteria, namely, the ability of the data extractors to understand other languages.

We included only studies in which arthroscopy was regarded as the standard of reference. This procedure has always been the reference standard for the diagnosis of internal derangements of the knee, against which alternative diagnostic modalities should be compared. However, the use of arthroscopy alone as the reference standard has been criticized because some parts of the joint cannot be brought into view properly. The posterior horn of the medial meniscus is an especially difficult area to visualize, and the arthroscopic diagnosis of meniscal tears in this region is often assigned on the basis of probing rather than visualizing the meniscus. Quinn and Brown (21) retrospectively analyzed the arthroscopic videotapes of false-positive MR imaging results and found that the suspected area of the meniscus was never visualized in these cases. Therefore, false-negative findings at arthroscopy could potentially account for many false-positive MR imaging results. Likewise, the PCL is not usually visualized during arthroscopy if the ACL is intact, and in this case, physical examination is often performed with the patient anesthetized to demonstrate a rupture of the PCL. As a result, arthroscopy is ideally performed with knowledge of the findings from the preceding MR examination.

We first summarized and combined the data in a pooled weighted analysis performed separately for each lesion. This type of analysis has certain limitations, since it does not take into account the differences in positivity criteria that were used in the studies. For the lesions that we included in this meta-analysis, however, we found that articles were highly comparable with regard to the definition of disease at MR imaging. In all included articles, the meniscus was evaluated and graded by using the original or slightly modified criteria that were introduced previously by Reicher et al (22) or Crues et al (23). On the basis of these criteria, a meniscal tear is diagnosed if an intrameniscal area of high signal intensity extends to the articular surface, whereas an area of high signal intensity that does not reach the surface is considered degenerative. Since degenerative changes of the menisci are common and usually asymptomatic findings in patients after the 3rd decade (24), our interest was focused on meniscal tears only. With regard to the status of the cruciate ligaments, a complete tear was usually defined as nonvisualization of the ligament in its expected position or high signal intensity in the course of the ligament in the absence of intact margins. Even though positivity criteria were largely the same among studies, we considered a summary ROC analysis appropriate, since, in our view, there will always be implicit variations in interpretation in different settings and by different radiologists.

The results of our meta-analysis stress the importance of adequate study design when conducting a study on diagnostic performance. Especially, the presence of verification bias was found to dramatically increase the diagnostic performance for PCL tears. Also, blinding the performer of the reference test (arthroscopy) to the findings of the test under evaluation (MR imaging) was predictive in the final model for medial meniscal tears. In every included study, the radiologist who interpreted the MR images was blinded to the arthroscopy result. This reflects clinical practice, in which MR imaging is usually and ideally performed prior to arthroscopy. Therefore, the effect of blinding the MR image interpreter to the result of the reference test was not assessed in this meta-analysis. In addition, mean age of the patient population was a statistically significant predictor for various lesions. This finding implies that the age distribution of the patient population should be described sufficiently in diagnostic performance studies.

The effect of publication year on the diagnostic performance probably reflects the progress in MR imaging technique over the past 15 years with regard to the MR imagers and the sequences used. Whereas in the early years of MR imaging, only spin-echo and gradient-echo sequences were performed, nowadays the choice of sequences that can be incorporated in the imaging protocol has expanded, with various types of fat-suppression techniques and three-dimensional acquisitions that enable reconstructions in every plane. Moreover, the influence of publication year could represent a learning process over the years among interpreters of MR images of the knee.

Higher magnetic field strength of the MR imager improved the diagnostic performance for ACL tears, and a modest effect was demonstrated when all lesions were combined in one model. In the pooled weighted analyses per lesion for different magnetic field strength categories, a trend was observed toward better diagnostic performance for higher magnetic field strengths, but this effect was far from significant, probably because too few studies were available per category of magnetic field strength. We can only speculate whether the effect would have been statistically significant if more existing studies would have fulfilled our inclusion criteria. Except for the ACL, however, no effect was demonstrated in the summary ROC analysis per lesion, which is a more powerful technique for assessing predictors of diagnostic performance. Only for PCL tears was the number of sequences a predictor in the final multivariate model. Unexpectedly, this effect was negative with increasing numbers of sequences. The status of the PCL can usually be assessed well on the basis of a few images in the sagittal plane.

We conclude that MR imaging is highly accurate in the diagnosis of tears of the menisci and cruciate ligaments. MR imaging is an appropriate screening tool for therapeutic arthroscopy, making diagnostic arthroscopy unnecessary in most patients. Diagnostic performance of MR imaging differs significantly for menisci and cruciate ligaments and for the medial and lateral meniscus and is influenced by characteristics of study design, patient age, and magnetic field strength. The effect of magnetic field strength, however, was only statistically significant for ACL tears. The influence of study design characteristics and patient age should be taken into consideration whenever a diagnostic performance study on MR imaging of the knee is designed or reported.

References

1. Rangger C, Klestil T, Kathrein A, Inderster A, Hamid L. Influence of magnetic resonance imaging on indications for arthroscopy of the knee. Clin Orthop 1996; 330: 133-142.
2. Rappeport ED, Mehta S, Wieslander SB, Lausten GS, Thomsen HS. MR imaging before arthroscopy in knee joint disorders? Acta Radiol 1996; 37:602-609.
3. Mackenzie R, Palmer CR, Lomas DJ, Dixon AK. Magnetic resonance imaging of the knee: diagnostic performance studies. Clin Radiol 1996; 51:251-257.
4. Irwig L, Tosteson AN, Gatsonis C, et al. Guidelines for meta-analyses evaluating diagnostic tests. Ann Intern Med 1994; 120:667-676.
5. Berlin JA. Does blinding of readers affect the results of meta-analyses? University of Pennsylvania Meta-analysis Blinding Study Group. Lancet 1997; 350:185-186.
6. Begg CB, Greenes RA. Assessment of diagnostic tests when disease verification is subject to selection bias. Biometrics 1983; 39:207-215.
7. Littenberg B, Moses LE. Estimating diagnostic accuracy from multiple conflicting reports: a new meta-analytic method. Med Decis Making 1993; 13:313-321.
8. Begg CB, Berlin JA. Publication bias and dissemination of clinical research. J Natl Cancer Inst 1989; 81:107-115.
9. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997; 315:629-634.
10. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986; 7:177-188.
11. Fleiss JL. The statistical basis of meta-analysis. Stat Methods Med Res 1993; 2:121-145.
12. Moses LE, Shapiro D, Littenberg B. Combining independent studies of a diagnostic test into a summary ROC curve: data-analytic approaches and some additional considerations. Stat Med 1993; 12:1293- 1316.

13. De Vries SO, Hunink MG, Polak JF. Summary receiver operating characteristic curves as a technique for meta-analysis of the diagnostic performance of duplex ultrasonography in peripheral arterial disease. Acad Radiol 1996; 3:361-369.
14. Irwig L, Macaskill P, Glasziou P, Fahey M. Meta-analytic methods for diagnostic test accuracy. J Clin Epidemiol 1995; 48:119- 130.
15. Vahey TN, Broome DR, Kayes KJ, Shelbourne KD. Acute and chronic tears of the anterior cruciate ligament: differential features at MR imaging. Radiology 1991; 181:251-253.
16. Warwick DJ, Cavanagh P, Bell M, Marsh CH. Influence of magnetic resonance imaging on a knee arthroscopy waiting list. Injury 1993; 24:380-382.
17. Munk B, Madsen F, Lundorf E, et al. Clinical magnetic resonance imaging and arthroscopic findings in knees: a comparative prospective study of meniscus anterior cruciate ligament and cartilage lesions. Arthroscopy 1998; 14:171-175.
18. Kawahara Y, Uetani M, Nakahara N, et al. Fast spin-echo MR of the articular cartilage in the osteoarthrotic knee: correlation of MR and arthroscopic findings. Acta Radiol 1998; 39:120-125.
19. Mori R, Ochi M, Sakai Y, Adachi N, Uchio Y. Clinical significance of magnetic resonance imaging (MRI) for focal chondral lesions. Magn Reson Imaging 1999; 17: 1135-1140.
20. Moher D, Pham B, Klassen TP, et al. What contributions do languages other than English make on the results of meta-analyses? J Clin Epidemiol 2000; 53:964-972.
21. Quinn SF, Brown TF. Meniscal tears diagnosed with MR imaging versus arthroscopy: how reliable a standard is arthroscopy? Radiology 1991; 181:843-847.
22. Reicher MA, Hartzman S, Duckwiler GR, Bassett LW, Anderson LJ, Gold RH. Meniscal injuries: detection using MR imaging. Radiology 1986; 159:753-757.
23. Crues JV III, Mink J, Levy TL, Lotysch M, Stoller DW. Meniscal tears of the knee: accuracy of MR imaging. Radiology 1987; 164:445-448.
24. Stoller DW, Martin C, Crues JV III, Kaplan L, Mink JH. Meniscal tears: pathologic correlation with MR imaging. Radiology 1987; 163:731-735.
25. Boeree NR, Watkinson AF, Ackroyd CE, Johnson C. Magnetic resonance imaging of meniscal and cruciate injuries of the knee. J Bone Joint Surg Br 1991; 73:452- 457.
26. Fischer SP, Fox JM, Del Pizzo W, Friedman MJ, Snyder SJ, Ferkel RD. Accuracy of diagnoses from magnetic resonance imaging of the knee: a multi-center analysis of one thousand and fourteen patients. J Bone Joint Surg Am 1991; 73:2- 10.
27. Kelly MA, Flock TJ, Kimmel JA, et al. MR imaging of the knee: clarification of its role. Arthroscopy 1991; 7:78-85.
28. Niitsu M, Anno I, Fukubayashi T, Shimojo H, Kuno S, Akisada M. Tears of cruciate ligaments and menisci: evaluation with cine MR imaging. Radiology 1991; 178:859-864.
29. Araki Y, Ootani F, Tsukaguchi I, et al. MR diagnosis of meniscal tears of the knee: value of axial three-dimensional Fourier transformation GRASS images. AJR Am J Roentgenol 1992; 158:587-590.

30. Gluckert K, Kladny B, Blank-Schal A, Hofmann G. MRI of the knee joint with a 3-D gradient echo sequence: equivalent to diagnostic arthroscopy? Arch Orthop Trauma Surg 1992; 112:5-14.
31. Grevitt MP, Pool CJ, Bodley RN, Savage PE. Magnetic resonance imaging of the knee: initial experience in a district general hospital. Injury 1992; 23:410-412.
32. Gross ML, Grover JS, Bassett LW, Seeger LL, Finerman GA. Magnetic resonance imaging of the posterior cruciate ligament: clinical use to improve diagnostic accuracy. Am J Sports Med 1992; 20:732- 737.
33. Heron CW, Calvert PT. Three-dimensional gradient-echo MR imaging of the knee: comparison with arthroscopy in 100 patients. Radiology 1992; 183:839- 844.
34. Barnett MJ. MR diagnosis of internal derangements of the knee: effect of field strength on efficacy. AJR Am J Roentgenol 1993; 161:115-118.
35. De Smet AA, Norris MA, Yandow DR, Quintana FA, Graf BK, Keene JS. MR diagnosis of meniscal tears of the knee: importance of high signal in the meniscus that extends to the surface. AJR Am J Roentgenol 1993; 161:101-107.
36. Chan WP, Peterfy C, Fritz RC, Genant HK. MR diagnosis of complete tears of the anterior cruciate ligament of the knee: importance of anterior subluxation of the tibia. AJR Am J Roentgenol 1994; 162: 355-360.
37. Gentili A, Seeger LL, Yao L, Do HM. Anterior cruciate ligament tear: indirect signs at MR imaging. Radiology 1994; 193:835-840.
38. Kinnunen J, Bondestam S, Kivioja A, et al. Diagnostic performance of low field MRI in acute knee injuries. Magn Reson Imaging 1994; 12:1155-1160.
39. LaPrade RF, Burnett QM, Veenstra MA, Hodgman CG. The prevalence of abnormal magnetic resonance imaging findings in asymptomatic knees: with correlation of magnetic resonance imaging to arthroscopic findings in symptomatic knees. Am J Sports Med 1994; 22:739- 745.
40. Chen MC, Shih TT, Jiang CC, Su CT, Huang KM. MRI of meniscus and cruciate ligament tears correlated with arthroscopy. J Formos Med Assoc 1995; 94:605- 611.
41. Justice WW, Quinn SF. Error patterns in the MR imaging evaluation of menisci of the knee. Radiology 1995; 196:617-621.
42. Lerman JE, Gray DS, Schweitzer ME, Bartolozzi A. MR evaluation of the anterior cruciate ligament: value of axial images. J Comput Assist Tomogr 1995; 19:604- 607.
43. Barry KP, Mesgarzadeh M, Triolo J, Moyer R, Tehranzadeh J, Bonakdarpour A. Accuracy of MRI patterns in evaluating anterior cruciate ligament tears. Skeletal Radiol 1996; 25:365-370.
44. Lundberg M, Odensten M, Thuomas KA, Messner K. The diagnostic validity of magnetic resonance imaging in acute knee injuries with hemarthrosis: a single-blinded evaluation in 69 patients using high-field MRI before arthroscopy. Int J Sports Med 1996; 17:218-222.
45. Riel KA, Reinisch M, Kersting-Sommerhoff B, Hof N, Merl T. 0.2-T magnetic resonance imaging of internal lesions of the knee joint: a prospective arthroscopically controlled clinical study. Knee Surg Sports Traumatol Arthrosc 1999; 7:37-41.

46. Bui-Mansfield LT, Youngberg RA, Warme W, Pitcher JD, Nguyen PL. Potential cost savings of MR imaging obtained before arthroscopy of the knee: evaluation of 50 consecutive patients. AJR Am J Roentgenol 1997; 168:913-918.
47. Cheung LP, Li KC, Hollett MD, Bergman AG, Herfkens RJ. Meniscal tears of the knee: accuracy of detection with fast spin-echo MR imaging and arthroscopic correlation in 293 patients. Radiology 1997; 203:508-512.
48. Franklin PD, Lemon RA, Barden HS. Accuracy of imaging the menisci on an in-office, dedicated, magnetic resonance imaging extremity system. Am J Sports Med 1997; 25:382-388.
49. Rappeport ED, Wieslander SB, Stephensen S, Lausten GS, Thomsen HS. MRI preferable to diagnostic arthroscopy in knee joint injuries: a double-blind comparison of 47 patients. Acta Orthop Scand 1997; 68:277- 281.
50. Weinstabl R, Muellner T, Vecsei V, Kainberger F, Kramer M. Economic considerations for the diagnosis and therapy of meniscal lesions: can magnetic resonance imaging help reduce the expense? World J Surg 1997; 21:363-368.
51. Ha TP, Li KC, Beaulieu CF, et al. Anterior cruciate ligament injury: fast spin-echo MR imaging with arthroscopic correlation in 217 examinations. AJR Am J Roentgenol 1998; 170:1215-1219.
52. Cotten A, Delfaut E, Demondion X, et al. MR imaging of the knee at 0.2 and 1.5 T: correlation with surgery. AJR Am J Roentgenol 2000; 174:1093-1097.
53. Elvenes J, Jerome CP, Reikeras O, Johansen O. Magnetic resonance imaging as a screening procedure to avoid arthroscopy for meniscal tears. Arch Orthop Trauma Surg 2000; 120:14-16.

Comparison of Short and Standard Protocols in Dedicated Extremity MRI of Wrist, Knee, and Ankle Injury

Jeroen J. Nikken M.D.[1,2], Abida Z. Ginai M.D. Ph.D.[1], Gabriel P. Krestin M.D. Ph.D.[1], M.G. Myriam Hunink M.D. Ph.D.[1,2,3]

From the Program for the Assessment of Radiological Technology (ART program), Department of Radiology[1] and Department of Epidemiology and Biostatistics[2], Erasmus University Medical Center Rotterdam, the Netherlands and the Department of Health Policy and Management[3], Harvard School of Public Health, Boston, MA.

Abstract
An important hindrance for the routine use of MRI in the initial evaluation of acute peripheral joint trauma is the high costs as well as the long duration of the examination. The aim of this study was to develop a short MRI protocol for the initial examination of peripheral joint trauma on a dedicated extremity MRI system and to compare this protocol with a standard protocol. Various parameters of sequences used for MRI of joints were adapted in order to decrease scanning time while preserving an acceptable diagnostic quality. The four sequences that appeared most suitable for the intended purpose (GE-T1w, SE-T1w, Turbo Multi Echo and STIR) were compared to standard MRI examination in 80 patients with peripheral joint trauma. The short MRI protocol as a screening examination for the presence or absence of pathology showed a sensitivity of 0.99 and a specificity of 0.90. Fracture and bone marrow edema could be demonstrated accurately with a sensitivity of 1.00 and 0.91 respectively and a specificity of 1.00 for both. The sensitivity and specificity for anterior cruciate ligament rupture was 0.78 and 0.90 respectively. For meniscus tears the specificity was high (medial 0.95, lateral 0.94), but the sensitivity was lower (medial 0.73, lateral 0.58). In conclusion, the short MRI examination cannot replace the standard examination for the comprehensive examination of wrist, knee, and ankle for all indications, however, it may be suitable as a screening tool in the initial investigation of patients with acute peripheral joint trauma.

Introduction
The role of MR imaging in the diagnosis and management of acute peripheral joint trauma is limited; physical examination and conventional radiographs are the main diagnostic modalities. In acute trauma it is often difficult to assess the extent of soft tissue injury by physical examination alone, as pain and swelling often prevent a proper physical examination. Conventional radiographs show fractures but provide little information on soft tissue injury. If MR imaging would be readily available for patients with acute joint injury, this diagnostic gap may be filled. The introduction of low field dedicated extremity MRI systems with their relatively low purchase price and maintenance costs has created the possibility of using MRI as a routine examination in patients with trauma of peripheral joints.

As far as we know, the routine use of MRI as initial examination in peripheral joint trauma has not yet been addressed in the literature. The use of MRI as an additional imaging modality, however, has been previously evaluated (1-6). Several studies show comparable results using extended protocols on high field and low field MR systems regarding the detection of internal derangement of the knee, using arthroscopy as reference standard (7-10). A prerequisite for a routine and cost-effective use of MRI in patients with acute joint injury is a short scanning time.

The purpose of this study was to optimize the scanning parameters of a short MRI protocol on a low field dedicated extremity MRI system. Our aim was to define a scanning protocol that would take approximately 10 minutes and that would be of sufficient quality to be used as an initial screening examination in peripheral joint trauma.

Materials and Methods

The first part of our study consisted of the development of a short scanning protocol. On a low field (0.2T) dedicated extremity MRI system (Artoscan M, Esaote, Genoa, Italy) we adapted the parameters of frequently used sequences in MR imaging for the evaluation of peripheral joints with the objective of shortening the scanning time considerably.

The following adaptations were applied in order to achieve short scanning times:

- short repetition time
- reduced number of acquisitions
- half Fourier transformation
- half echo acquisition
- rectangular Field of View (FOV)
- limited matrix size

Table 1 shows the sequences that were evaluated in 10 healthy volunteers. In shortening the repetition times of Gradient Echo (GE) and T1-weighted Spin Echo (SE) we were limited by the minimal repetition time (TR) necessary for multislice imaging. The Turbo Multi Echo (TME) sequence uses two echoes; the first echo has a fixed effective echo time of 38 ms, the second echo has a fixed effective echo time of 90 ms. We used a TR of 2400 ms to obtain sufficient T2 weighting. To find the optimal FOV we systematically varied the phase encoding (time dependent) direction of the FOV from 10 cm (minimum) to 20 cm (maximum) with incremental steps of 1 cm. The smallest FOV with no disturbing fold-over artefacts in the region of interest was chosen. We evaluated the effect of limiting the matrix size while retaining an acceptable spatial resolution by varying the matrix size from 256x192 to 128x192 in 16 steps. The smallest matrix that subjectively still provided

enough detail was chosen. The images were evaluated by two experienced radiologists, who made the choice of optimal FOV and matrix in consensus.
The second part of the study comprised the comparison of the short scanning protocols, developed and selected in the first part, with the standard protocols. We examined 80 patients (32 female, 48 male, age range 15 - 78 year, mean age 33 year) with acute joint injury, referred to us by traumatologist or orthopaedic surgeon. These consisted of 48 knees, 23 ankles, and 9 wrists. For the wrist we used sagittal T1-w SE, axial TME, coronal GE and coronal Short Tau Inversion Recovery (STIR); for the knee we used sagittal T1-w SE, coronal TME, coronal GE and sagittal STIR; for the ankle we used axial TME, sagittal T1-w SE, coronal GE and coronal STIR. Standard MRI examination on the same MRI system, performed immediately following the short protocol, was used as reference test. The standard MRI protocol for the ankle and knee consisted of T1-w half SE, T2-w turbo SE, TME, GE and STIR sequences; for the wrist we used TME, 3-D T1-w SE, STIR and GE sequences.
The short and the standard MRI examinations were assessed separately by an experienced musculoskeletal radiologist, with an intervening interval of on average 40 days. Meniscus tear was defined as a high intensity linear signal in the meniscus reaching the articular surface. Bone marrow edema was defined as a high signal intensity in the bone on the STIR sequence. For the diagnosis of fracture the lesion had to be visible both on STIR as well as on one of the other sequences. Ligament rupture was defined as a discontinuity of the ligament on a T1-w SE or T2-w SE sequence. Additional signs of anterior cruciate ligament (ACL) rupture were horizontal or vertical orientation, non-visualisation or diffuse increase in signal on T2-w images, and enlargement of the ligament (11). Partial ACL rupture was defined as high intensity signal in the ACL with visibly intact fibers passing the region of high signal intensity.
Using the standard MRI as reference test, we determined sensitivity, specificity, and the positive and negative predictive values of the short protocol for various common abnormalities. Sensitivity was calculated as the proportion of patients with a particular abnormality on the standard MRI series, that demonstrated the same abnormality on the short MRI series. Specificity was calculated as the proportion of subjects without the abnormality on the standard MRI series that did not demonstrate the abnormality on the short MRI series. The positive predictive value was calculated as the post-positive test probability of the abnormality and the negative predictive value was calculated as the post-negative test probability of absence of the abnormality.

The study was approved by the institutional review board, all subjects were given both oral and written explanation of the study and procedures and they all gave written informed consent.

Results

Table 1 presents the minimal acceptable FOV and matrix size required for the sequences under study, as well as their mean duration. A FOV of 14 cm in the phase encoding direction was always enough to prevent disturbing aliasing in the region of interest. The minimal acceptable matrix size varied considerably between the various sequences. The sequences that use partial filling of k-space required a larger matrix size than the other sequences. Based on clinical applicability and mean scanning time we selected the T1-w SE with half-Fourier transformation, T1-w GE, TME and STIR sequences for use in the short protocol. The total scanning time using this short protocol, including calibration and calculation time, was 12 minutes.

Table 1. Parameters of the short sequences under evaluation

	Phase encoding direction of FOV*(cm)	Matrix size	TR (ms)	TE (ms)	Mean scanning time (min:s)
T1-w SE	14	192*144	550	24	1:39
T1-w SE-half Fourier	13	192*192	550	24	1:03
T1-w SE-half Echo	13	192*192	550	14	1:18
T2-w turbo SE	14	192*144	2400	90	2:43
Turbo Multi Echo	14	192*128	2400	38 and 90	2:31
Gradient Echo**	14	192*160	350	38	0:45
STIR***	13	192*128	1000	24	1:30

FOV : field of view, TR : repetition time, TE : echo time, SE : spin echo
For all sequences: number of excitations (NEX)=1
* Frequency encoding direction: 20 cm for all sequences
** Flip angle: 75 degree
*** STIR : time to inversion: 80 ms

Table 2 presents the test characteristics of findings in the 80 studied patients. Evaluating the presence or absence of any pathology, we found a sensitivity of 0.99 and a specificity of 0.90. Fluid was frequently present in the joint after trauma. The short protocol had a sensitivity of 0.96 and a specificity of 0.90 for fluid detection in the joint. All fractures were diagnosed correctly with no false positive or false negative outcomes. Two of the fractures were radiologically occult; one was a fracture of the tibia plateau and the other a fracture of the patella. The STIR sequence was very sensitive for the diagnosis of bone marrow edema and it showed 20 out of 22 cases of bone marrow edema without any false positive findings. The two cases in which bone marrow edema was not seen using the short protocol had in fact only very slight changes of signal intensity on the standard MRI images.

Table 2. Test characteristics of the short protocol compared to the standard protocol as reference test for various abnormalities.

	Sensitivity	Specificity	PPV	NPV
Any abnormality	69/70 (0.99)	9/10 (0.90)	0.99	0.90
Fluid in the joint	48/50 (0.96)	27/30 (0.90)	0.94	0.93
Fracture	6/6 (1.00)	74/74 (1.00)	1.00	1.00
Bone edema	20/22 (0.91)	58/58 (1.00)	1.00	0.97
Medial meniscus tear	8/11 (0.73)	35/37 (0.95)	0.80	0.92
Lateral meniscus tear	7/12 (0.58)	34/36 (0.94)	0.78	0.87
ACL rupture	7/9 (0.78)	35/39 (0.90)	0.64	0.95
ACL rupture or partial rupture	21/25 (0.84)	22/23 (0.96)	0.95	0.85
ATFL rupture	6/10 (0.60)	13/13 (1.00)	1.00	0.76
Syndesmosis rupture	3/4 (0.75)	19/19 (1.00)	1.00	0.95

PPV : positive predictive value, i.e. post-positive test probability of the abnormality
NPV : negative predictive value, i.e. post-negative test probability of absence of the abnormality
ACL : anterior cruciate ligament
ATFL : anterior talofibular ligament

Sensitivity was moderate for medial meniscus tear (73%) and was poor for lateral meniscus tear (58%). Specificity, however, was high for both medial and lateral meniscus tears (95% and 94%). For rupture of the anterior talofibular ligament we found a low sensitivity of 60%, but specificity was 100%. For anterior cruciate ligament rupture the sensitivity was 78% and specificity 90 %. Of the 9 ACL ruptures 7 were diagnosed correctly and two were classified as partial rupture. The four false positive diagnoses of ACL rupture were all found to be partial ruptures.

Discussion

Injuries of ankle, knee, and wrist are a very common reason for patients to visit an emergency department. The relevance for health care policy of these injuries is illustrated by the fact that 527 million Euro were spent on the management of accidental falls (excluding motor vehicle accidents) in the Netherlands in the year 1994 (12). Ankle, knee, and wrist trauma represent a substantial proportion of these accidents. The costs to society include not only medical costs but also the costs due to loss of productivity. If the time to diagnosis can be shortened by using MRI in the initial evaluation, this may lead to earlier and more adequate treatment and hence earlier recovery. MRI can also be cost saving in patients without structural damage. Uncertainty of the diagnosis is sometimes a reason for temporary treatment measures. A normal MRI can obviate the need for these measures as well as for follow-up visits and additional investigations. Although performing a short MRI routinely for peripheral joint trauma may increase the cost of the initial workup, such a strategy may decrease the overall costs to society. The feasibility of using MRI in the initial workup is greatly enhanced by the development of relatively inexpensive low field dedicated extremity MRI systems. The reason for undertaking the present study was to provide a short MRI scanning protocol for the purpose of initial screening in peripheral joint trauma and to compare it with the standard extended MRI protocol.

The choice of the scanning parameters is a trade-off between image quality and scanning time. We determined, subjectively, the minimum required spatial resolution and contrast that provided images of sufficient quality. The total scanning time of the short protocol was between 10 and 12 minutes, depending on the size of the joint, which we think will be acceptable for routine use in acutely injured joints.

The use of a short MRI protocol as a screening test will only be effective if it has the ability to predict the presence vs. absence of pathology. Our short protocol did indeed provide a high positive predictive value (PPV) and a reasonable negative

predictive value (NPV) for the detection of "any" abnormality. The most common abnormality found was fluid in the joint. Using the short protocol the presence of fluid in the joint could be detected fairly accurately, but it was, however, not a strong indicator of structural damage.
Although the number of fractures in our patient group was relatively low, the results (no false positive or false negative findings) indicate that the short protocol is likely to be suitable for the detection of fractures. MRI has been shown to be very suitable to detect radiologically occult fractures; especially the STIR sequence is very sensitive. (13-18). Our findings support these earlier reported results.
Several studies report good performance of low field MRI in the detection of meniscus lesions (7-10). Using the short protocol it is likely that some meniscus tears will be missed, as indicated by the moderate sensitivity. Most of the missed meniscus tears, however, appeared to be small tears and not bucket handle tears, flap tears, or separations. The treatment policy for these small tears will generally be conservative, since it is unlikely that these small tears will cause complaints in the long term. In fact, small meniscus tears may heal without any specific treatment (19, 20).
The short protocol allows detection of ACL lesions (i.e. both rupture and partial rupture of the ACL). If the ACL lesion was accompanied by typical bone marrow edema in the lateral femur condyl and the posterolateral tibia plateau the distinction between partial rupture and total rupture was always clear. However, if no bone marrow edema was present the distinction between partial and total ruptures was poor. In these cases an additional standard sequence suitable for depicting ACL lesions will be necessary.
For anterior talofibular ligament lesions the predictive value of a negative short MRI examination is only moderate. In general, however, early detection of anterior talofibular ligament rupture is not considered very important, as it does not change the short-term treatment of the patient. The value of the short MRI examination in ankle trauma for the detection of syndesmosis rupture is, however, promising although it must be noted that there were only four patients with this type of lesion in the study group.

It must be noted that in this study the standard MRI was the reference test. Because the reference test itself also yields false positive and false negative findings, we may have overestimated the accuracy of the short protocol to some extent. We did not use arthroscopy as reference test, because only a few patients actually underwent arthroscopy and subjecting all subjects to arthroscopy would not have been ethically justified. Apart from ethical considerations, the value of arthroscopy

as reference test can be questioned, since most lesions were not intra-articular and would therefore not have been seen on arthroscopy.
Implementing an initial short MRI examination could potentially lead to reimbursement problems if health insurance companies and HMOs are not prepared to pay for a second extensive MRI examination for the same clinical problem. As with all new applications of technology, we would need to demonstrate the cost-effectiveness of the strategy we are advocating to convince payers to reimburse both examinations. Using a short protocol instead of the standard protocol implies accepting some decrease in diagnostic quality. In spite of this limitation the application of the short protocol may be valuable if the additional diagnostic information leads to a change in treatment and subsequently to an earlier recovery.

We expect that using a short MRI series in the initial examination of peripheral joint trauma will obviate the need for temporary treatment, follow-up, and additional examinations in patients with normal findings on MRI. It may also save costs through earlier return to normal productivity if in patients with positive MRI findings early diagnosis leads to a change in treatment and hence to a faster recovery. Whether an initial short MRI examination in all patients with peripheral joint injury would be cost-effective compared to the present strategy subject of study in following chapters.
In conclusion, the short MRI protocol developed on a low field MRI system, taking only about 12 minutes, may be a suitable routine investigation in patients with acute trauma of wrist, knee or ankle.

References

1. Polly DW, Jr., Callaghan JJ, Sikes RA, McCabe JM, McMahon K, Savory CG (1988) The accuracy of selective magnetic resonance imaging compared with the findings of arthroscopy of the knee. J Bone Joint Surg [Am]70(2):192-8.
2. Mink JH, Levy T, Crues JVd. (1988) Tears of the anterior cruciate ligament and menisci of the knee: MR imaging evaluation. Radiology 167(3):769-74.
3. Fischer SP, Fox JM, Del Pizzo W, Friedman MJ, Snyder SJ, Ferkel RD (1991) Accuracy of diagnoses from magnetic resonance imaging of the knee. A multi-center analysis of one thousand and fourteen patients. J Bone Joint Surg [Am]73(1):2-10.
4. Justice WW, Quinn SF (1995) Error patterns in the MR imaging evaluation of menisci of the knee. Radiology 196(3):617-21.
5. Boeree NR, Watkinson AF, Ackroyd CE, Johnson C (1991) Magnetic resonance imaging of meniscal and cruciate injuries of the knee. J Bone Joint Surg Br 73(3):452-7.
6. De Smet AA, Tuite MJ, Norris MA, Swan JS (1994) MR diagnosis of meniscal tears: analysis of causes of errors [see comments]. AJR Am J Roentgenol 163(6):1419-23.

7. Barnett MJ (1993) MR diagnosis of internal derangements of the knee: effect of field strength on efficacy. AJR Am J Roentgenol 161(1):115-8.
8. Glashow JL, Katz R, Schneider M, Scott WN (1989) Double-blind assessment of the value of magnetic resonance imaging in the diagnosis of anterior cruciate and meniscal lesions. J Bone Joint Surg [Am] 71(1):113-9.
9. Mandelbaum BR, Finerman GA, Reicher MA, Hartzman S, Bassett LW, Gold RH, et al (1986) Magnetic resonance imaging as a tool for evaluation of traumatic knee injuries. Anatomical and pathoanatomical correlations. Am J Sports Med 14(5):361-70.
10. Franklin PD, Lemon RA, Barden HS (1997) Accuracy of imaging the menisci on an in-office, dedicated, magnetic resonance imaging extremity system. Am J Sports Med 25(3):382-8.
11. Barry KP, Mesgarzadeh M, Triolo J, Moyer R, Tehranzadeh J, Bonakdarpour A (1996) Accuracy of MRI patterns in evaluating anterior cruciate ligament tears. Skeletal Radiol 25(4):365-70.
12. Polder JJ (1997) Kosten van ziekten in Nederland 1994. Rotterdam: Instituut Maatschappelijke Gezondheidszorg
13. Schick S, Trattnig S, Gabler C, Kukla C, Gahleitner A, Kainberger F, et al (1999) Okkulte Handgelenksfrakturen: Feinfokusvergrosserungsrontgen versus MRT. Rofo Fortschr Geb Rontgenstr Neuen Bildgeb Verfahr 170(1):16-21.
14. Kitsis C, Taylor M, Chandey J, Smith R, Latham J, Turner S, et al (1998) Imaging the problem scaphoid. Injury 29(7):515-20.
15. Bretlau T, Christensen OM, Edstrom P, Thomsen HS, Lausten GS (1999) Diagnosis of scaphoid fracture and dedicated extremity MRI. Acta Orthop Scand 70(5):504-8.
16. Breitenseher MJ, Metz VM, Gilula LA, Gaebler C, Kukla C, Fleischmann D, et al (1997) Radiographically occult scaphoid fractures: value of MR imaging in detection [see comments]. Radiology 203(1):245-50.
17. Jerosch J, Castro WH, Lahm A, Assheuer J (1991) Okkulte ossare Begleitverletzungen bei Verletzungen des Kniegelenkes. Sportverletz Sportschaden 5(4):193-8.
18. Van Gelderen W, Gale RS, Steward AH (1998) Short tau inversion recovery magnetic resonance imaging in occult scaphoid injuries: effect on management. Australas Radiol 42(1):20-4.
19. Quinn SF, Brown TF (1991) Meniscal tears diagnosed with MR imaging versus arthroscopy: how reliable a standard is arthroscopy? Radiology 181(3):843-7.
20. LaPrade RF, Burnett QM, 2nd, Veenstra MA, Hodgman CG (1994) The prevalence of abnormal magnetic resonance imaging findings in asymptomatic knees. With correlation of magnetic resonance imaging to arthroscopic findings in symptomatic knees. Am J Sports Med 22(6):739-45.

Acute Wrist Trauma - the Value of A Short Dedicated Extremity MRI Examination in Predicting Subsequent Treatment

Jeroen J. Nikken, MD [1,2], Edwin H.G. Oei, MSc [1,2], Abida Z. Ginai, MD, PhD [1], Gabriel P. Krestin, MD, PhD [1], Jan A.N. Verhaar, MD, PhD [3], Arie B. van Vugt, MD, PhD [4], Myriam G.M. Hunink, MD, PhD [1,2]

From the Program for the Assessment of Radiological Technology (ART Program).
Department of Radiology [1], Department of Epidemiology and Biostatistics [2],
Department of Orthopaedic Surgery [3] and Department of Traumatology [4],
Erasmus MC, University Medical Center, Rotterdam. the Netherlands.

Submitted for publication.

Abstract

Purpose
To assess the value of a short MRI examination in addition to or instead of radiography in patients with acute wrist trauma in identifying patients who need additional treatment versus patients not requiring therapy, who can be safely discharged from further follow-up.

Methods and materials
Ninety patients with acute wrist trauma were randomized between radiography and radiography followed by a short MRI examination on a low field dedicated extremity MR system. Data were collected on sex, age, trauma mechanism, tenderness of the anatomical snuffbox, radiography result, MRI result, and treatment. Univariable and multivariable logistic regression analyses were used to create four models predicting the need for treatment.

Results
Thirty-six fractures were found, only one patient had a significant soft tissue lesion. In univariable analysis, age (Odds ratio (OR) 1.02, 95% confidence interval (CI) 1.00-1.05), snuffbox tenderness (OR 2.31, 95% CI 0.90-5.69), radiography result (OR 31.2, 95% CI 8.90-109), and positive MRI result vs no MRI made (OR 1.86, 95% CI 0.57-6.06) significantly predicted treatment. In the multivariable analysis, radiography result (OR 24.7, 95% CI 6.59-93.1) and positive MRI result (OR 6.28, 95% CI 1.27-31.0) contributed significantly in predicting the need for treatment. However, a negative MRI result had no significant predictive value (OR 0.87, 95% CI 0.20-3.8).

Conclusion
A short MRI examination on a low field MRI system following radiography in the initial evaluation of patients with acute wrist trauma has additional value in predicting subsequent treatment, but it doesn't add in the identification of patients who can be safely discharged from further follow-up.

Introduction

The standard diagnostic workup of acute wrist trauma consists of physical examination, often complemented by radiography. This diagnostic strategy demonstrates most fractures and luxations, but it provides little information on soft tissues like tendons, ligaments and the triangular fibrocartilage complex (TFCC). Another diagnostic problem in the injured wrist is the relatively high rate of occult scaphoid fractures (1-2). If an initially missed scaphoid fracture is not detected in time, the delay in immobilization considerably increases the risk of non-union (3) with potentially avascular necrosis as a result. A clinically suspected scaphoid fracture, not visible on initial radiography, is commonly immobilized and re-evaluated by means of repeat radiography. It may, however, take up to 6 weeks for a fracture to become visible on the radiograph (1). In this strategy about three out of four patients will be unnecessarily immobilized (4). Several studies show high accuracy of both high and low-field MRI in the detection of radiologically occult scaphoid fractures (1, 5-8).

Whereas a scaphoid fracture can be clinically suspected by tenderness of the anatomical snuffbox, a tear of the triangular fibrocartilage complex (TFCC) or a rupture of interosseous carpal ligaments cannot be detected by initial physical examination. Initial MRI examination may play a role in the detection of these lesions as well, although it's role will likely be more limited in this respect. Early MRI examination has been found useful in the detection of scapholunate ligament ruptures and TFCC tears (7), but others state that for the scapholunate ligament the specificity may be high, but the sensitivity is low (9-12).

The high costs of high field MRI systems have prevented a widespread use of MRI in acute wrist injury. However, the gradual rise of low field dedicated extremity MRI systems may have changed this situation, since these systems are considerably less expensive than high field MRI systems, both in acquisition as well as in utilization, and their development is progressing rapidly.

The purpose of this study was to evaluate the value of a short MRI examination, using a low-field dedicated system, in all patients presenting with acute wrist trauma in predicting the need for treatment after the initial visit. The potential benefit of an early MRI examination lies in identifying patients that do not need additional treatment and who can be safely discharged from further follow-up. Moreover, patients that do have significant wrist injury not visible on the radiograph may be diagnosed earlier if a short MRI examination were to be performed at the initial visit. Using this strategy costs may be saved on the one hand by a reduced use of medical resources and on the other hand by decreasing the time to final treatment and therefore potentially diminishing time off work, thereby reducing productivity loss.

Materials and Methods

Study design

The data for this study were obtained from a prospective randomized controlled trial (RCT) addressing the effectiveness of early MR scanning for trauma of the wrist. Patients were eligible for the study if they sustained a recent injury of the wrist (within 7 days of trauma) and were referred to the radiology department by a traumatologist or orthopaedic surgeon for radiography of the affected joint. All patients were informed about the study by leaflet and by oral explanation. Informed consent was obtained from all participating patients. The study was approved by the institutional review board.

The patients were randomized between plain radiography alone and radiography combined with MRI examination. Randomization was carried out by drawing from consecutively numbered sealed envelopes containing computer generated random assignments. Exclusion criteria were additional injury of head, back, thorax or abdomen, compound fracture of the wrist, pre-existing complaints of the same wrist and intoxication by alcohol or drugs. Patients were included from 8.00 h to 23.00 h, 7 days a week by research staff or radiology technologists on service.

Radiographs were taken in posteroanterior and lateral direction. If a scaphoid fracture was suspected a scaphoid X-ray series was performed, consisting of a posteroanterior view with the hand in ulnar deviation, a lateral view and two oblique views. Immediately after the radiography, an MRI was made on a 0.2 Tesla dedicated extremity MRI system (Artoscan M, Esaote Biomedica, Genoa, Italy). Our standard MRI protocol for the wrist takes about 30 minutes. This would be both too time consuming and too costly for application in all patients with acute wrist injury. Therefore we used a shortened MRI scanning protocol (table 1). This protocol was achieved by using only one excitation for each sequence, a rectangular field of view and by limiting the phase encoding direction of the matrix. Shortening of the duration of the examination comes at the cost of some loss in quality, but in a pilot study including 48 patients with recent knee trauma we found that the subjective image quality was acceptable. The average acquisition time was 5.34 minutes with a total examination time, including start-up of the MRI system and patient positioning, of approximately 15 minutes. The MRI was assessed immediately by an experienced musculoskeletal radiologist or a radiology resident without knowledge of the radiography result. The MRI result was reported to the physician, so the information was available for the decision on treatment or follow-up of the patient.

TABLE 1. Parameters of the short MRI protocol

	Plane	TR (ms)	TE (ms)	Matrix	Slice thickness (mm)	scan time (min:s)
(HF) SE T1-w	Sagittal	420-630	18-24	192 x 216	3.0	0:43
TME PD-w and T2-w	Axial	2310-2480	28-38/90	256 x 128	3.0	2:25
GE *	Coronal	360-420	15	192 x 152	3.0	0:45
STIR **	Coronal	830-1060	24-28	192 x 128	4.0	1:37

HF = half Fourier, SE = spin echo, TME = turbo multi echo, PD = proton density, GE = gradient echo, STIR = short-tau inversion recovery, TR = repetition time, TE = echo time.
For all sequences: field of view=200x140 mm, number of excitations = 1
* Flip angle : 75 degrees
** Inversion time : 80-85 ms

MRI examinations assessed by the radiology resident (e.g. during weekends and in the evenings) were reassessed by the musculoskeletal radiologist the next day, without knowledge of the radiography result, and if the interpretation was different, the treating physician was informed. The radiographs were assessed by the physician in the emergency ward initially, conform daily practice in our hospital, and the next day the radiographs were reassessed by the musculoskeletal radiologist in a plenary session with the traumatologist.

Outcome

As outcome measure we used additional treatment after initial presentation. We did not use initial treatment as outcome measure because almost all patients received some form of treatment at the first visit, if only a bandage in case of a contusion. By analyzing treatment after the initial hospital visit, we were able to assess if MRI could discriminate between patients who did not need follow-up versus those who needed additional treatment and follow-up.

Data collection

Data were collected on age, sex, side of trauma, trauma mechanism (direct or indirect trauma), tenderness of the anatomical snuffbox, radiography result, MRI result (traumatic injury visible, no traumatic injury visible or uncertain), and therapeutic procedures.

Data were obtained by direct assessment and by reviewing all information available from emergency department records, hospital records, outpatient clinic records, and the electronic hospital information system. Questionnaires were sent to all patients one week, six weeks, twelve weeks, and six months after the initial visit, to obtain information about treatment performed outside our hospital. In case questionnaires were not returned, we interviewed the patient by telephone, provided they could be reached within two attempts of calling them. The final diagnosis was obtained by reviewing all information of diagnostic imaging, follow-up and questionnaires. The follow-up period was 6 months. However, if the patient indicated on the questionnaire that he/she did not have daily complaints anymore the follow up was terminated, since no additional effects of the injury were to be expected.

Image interpretation

The radiograph and the MRI were scored as 'positive' if signs of possible recent traumatic injury were visible. For the radiograph these signs were fracture, luxation, epiphysiolysis, osteochondral lesion or scapholunate diastase, indicating rupture of the scapholunate ligament. The MRI was scored positive if one or more of the following lesions were present: fracture, luxation, epiphysiolysis, osteochondral lesion, ligament rupture or TFCC tear. Fluid in the joint or around tendons, hematoma, soft tissue edema, and bone marrow edema were considered negative findings, since specific treatment is generally not indicated for these findings. Of course, the presence of these findings will make the radiologist more cautious about the possible presence of a fracture or ligament tear.

Data analysis

The predictive value of age, sex, side, trauma mechanism, tenderness of the anatomical snuffbox, radiography result, and MRI result was analyzed using univariable logistic regression. For the MRI result the group of patients randomized to the 'no-MRI' group was used as reference to analyze the incremental value of the MRI information compared to absence of this information. Therefore the variable 'MRI result' was recoded into three different outcomes: 'positive or uncertain', 'negative', or 'no MRI performed'. In the regression analysis the MRI result was analyzed with the outcome 'no MRI performed' as reference. In this way the whole patient population could be used in the analysis of the MRI result.

To account for the fact that the radiograph and MRI contain overlapping information, we created an additional integrated 'imaging variable' containing both the radiograph and MRI result. This variable had five possible outcomes:

radiograph negative and MRI not performed, radiograph positive and MRI not performed, radiograph and MRI both negative, radiograph and MRI discrepant, and radiograph and MRI both positive. The outcome 'radiography negative and MRI not performed' was considered the reference group for the combined 'imaging variable' analysis. With this combined 'imaging variable' we took into account the diagnostic interaction between radiography and MRI.

A variable was included in the multivariable logistic regression analysis if the chi square p-value was less than 0.1 in the univariable analysis. We created four multivariable models predicting the for of additional treatment after the initial hospital visit. In model 1 only radiography was considered as imaging test; in model 2 radiography was omitted and MRI was used instead, model 3 contained both radiography and MRI as additive variables; and in model 4 the combined 'imaging variable' was used.

The Likelihood Ratio test was used in a stepwise backward approach to assess if the variable contributed significantly to the model. A p-value smaller than 0.05 indicated a significant contribution of the variable in the model. Hosmer and Lemeshow's Goodness of Fit Test was used to assess calibration of the models (13). A large p-value indicates good calibration of the model. The area under the Receiver Operating Charcteristic (ROC) curve was used to compare predictive performance across the models (14). The difference between the areas under the ROC curve of the different models was analyzed using the z-statistic for ROC curves and corrected for the fact that they were derived from the same cases (15). The correction factor was calculated using the Pearson product-moment correlation method.

Akaike's Information Criterion (AIC) was used to compare the predictive power of the models (16). AIC is derived from the chi-square statistic and takes into account the complexity and the accuracy of the model by correcting for the degrees of freedom. In this way direct comparison of models is possible. A value of AIC equal to or smaller than 0 indicates no predictive value of the model. The significance of a difference in AIC of nested models (e.g. the model with radiography vs the model with radiography and MRI) was tested using the likelihood ratio test. In non-nested models (e.g. the model with radiography vs the model with the 'combined imaging variable') this test of significance cannot be used. For these models a difference in AIC larger than 2 was considered significant. All analyses were performed with SPSS for Windows (release 10.0.7; SPSS Inc., Chicago, IL, USA).

Sample size

For each variable analyzed in the multivariable logistic regression analysis at least 10 patients with an event and at least 10 patients without an event are necessary (20). Because the data for this study were obtained from an RCT addressing the costs and effectiveness of early MR scanning of wrist trauma, the number of variables analyzed in the multivariable regression analysis was determined by the number of patients included in this RCT. Since 37 of the included patients needed additional therapy and 50 patients did not need additional therapy, we could analyze 4 variables in the multivariable regression analysis.

Results

From August 1999 to May 2001, 90 patients were included (37 female and 53 male, mean age 40.0 years, range 16.4-80.0 years). A flow diagram of subjects passing through the trial is listed in figure 1.

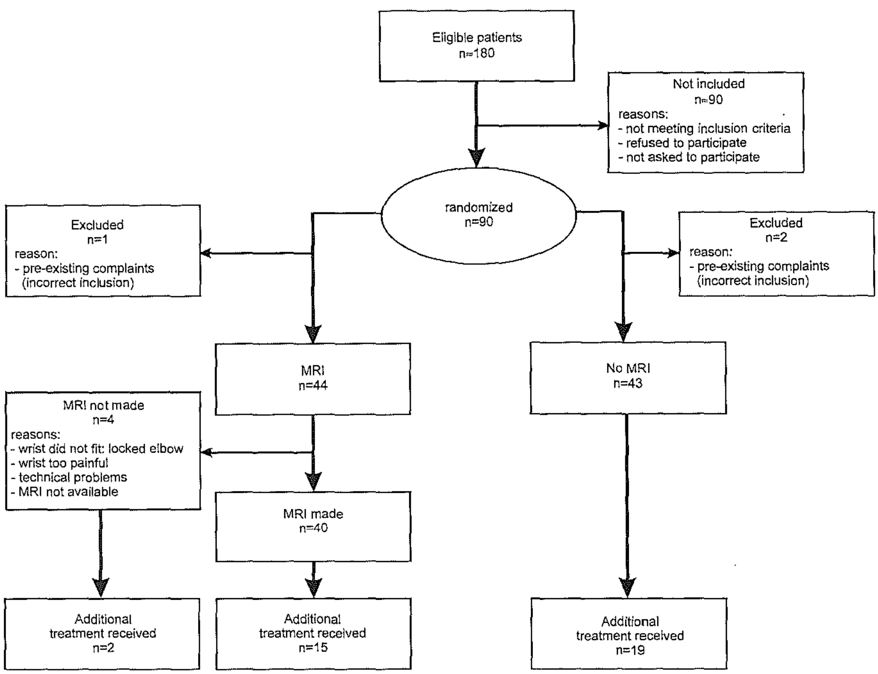


Figure 1. Flow diagram of subjects passing through the study.

Although inclusion was intended to be consecutive, only about half of all eligible patients was randomized. The reasons were refusal of patients to participate and patients that were inadvertently missed. After enrollment, three patients appeared to have pre-existing complaints of the wrist and they were subsequently excluded. Of the remaining 87 patients, 44 were allocated to the MRI strategy and 43 to the reference group. In 4 patients, allocated to the MRI strategy, the MRI was not made or not interpretable; in one case because of technical failure, in one case the examination was interrupted because of pain, in one case the MRI system was not available, and in one case the patient was not able to extend the elbow, so the joint could not reach the middle of the magnet bore.
The final diagnoses are listed in table 2. Two triangular fibrocartilage complex (TFCC) lesions were initially diagnosed on MRI, however, both patients were free of symptoms without specific treatment within a relatively short time (50 and 60 days). In these cases the MRI result may have been false positive or the TFCC lesion may have been asymptomatic.

TABLE 2. Final diagnosis in 86 patients with acute wrist injury. The number of patients in the experimental group and in the control group are noted.

Diagnosis	Experimental group	Control group	Total
Contusion	26	22	48
Fracture	16	20	36
Distal radius	2	8	10
Distal radius, radiographically occult (including 2 Torus fractures)	4		4
Distal radius + ulna	1	4	5
Distal radius + scaphoid fracture	0	1	1
Scaphoid	3	5	8
Scaphoid, radiographically occult	3		3
Triquetrum avulsion	2 *	2	4
Metacarpal	1 *		1
Rupture of extensor pollicis brevis tendon	1		1
No abnormality	1		1

* These fractures were only visible on the radiograph and not diagnosed on MRI. The triquetrum avulsions were not visible on MRI and the metacarpal fracture was out of the field of view

In one patient radiography showed a scapholunate diastasis of 3 mm, determined in the mid portion of the scapholunate joint, compared to a joint width of 1,5 mm in the capitolunate joint. MRI confirmed the scapholunate diastasis, as well as a suspected scapholunate ligament rupture. This patient, however, was free of symptoms within seven days after the trauma. It may well be possible that this patient had a (pre-existing) symptomless scapholunate ligament rupture, which has been described before (17).

Treatment of fractures consisted of splint, plastercast or rigid synthetic cast in most cases. In the four cases with an avulsion fracture of the triquetral bone and in three of four cases with occult distal radius fracture (all diagnosed on the initial MRI) no additional immobilization was applied after the first visit. One patient with occult distal radius fracture was treated with prolonged immobilization during follow-up. Five clinically suspected scaphoid fractures in the group with only radiography were immobilized temporarily with plastercast. Only one of these appeared to be a scaphoid fracture during follow-up. Another of the five patients clinically suspected of scaphoid fracture appeared to have a ruptured tendon of the extensor pollicis brevis muscle at the insertion on the proximal phalanx, which was not diagnosed on the MRI, since it was not within the field of view. Out of 12 patients with a scaphoid fracture, 9 had tenderness of the anatomical snuffbox, two did not have tenderness of the anatomical snuffbox and in one case the presence or absence of tenderness of the anatomical snuffbox was not reported. All patients with a scaphoid fracture that was occult on radiography, but visible on MRI (Figure 2), had tenderness of the anatomical snuffbox. No patients with a negative MRI result regarding scaphoid fracture were found to hav a fracture during follow-up. In seven cases of all fractures, immobilization was followed by physical therapy. Furthermore, physical therapy was given in four cases of contusion of the wrist. One patient with a distal radius and ulna fracture (Smith type) was operated six weeks after the trauma because conservative therapy had failed.

Regression analysis

Univariable analysis of the variables (Table 3) showed that age, tenderness of the anatomical snuffbox, radiography result and MRI result were significant predictors of additional treatment using a p-value threshold of 0.1. Sex and side did not significantly predict treatment.

Since there were only 36 end-points (36 patients treated) and given the low relative odds ratio for age, we chose not to use age in the multivariable model to avoid

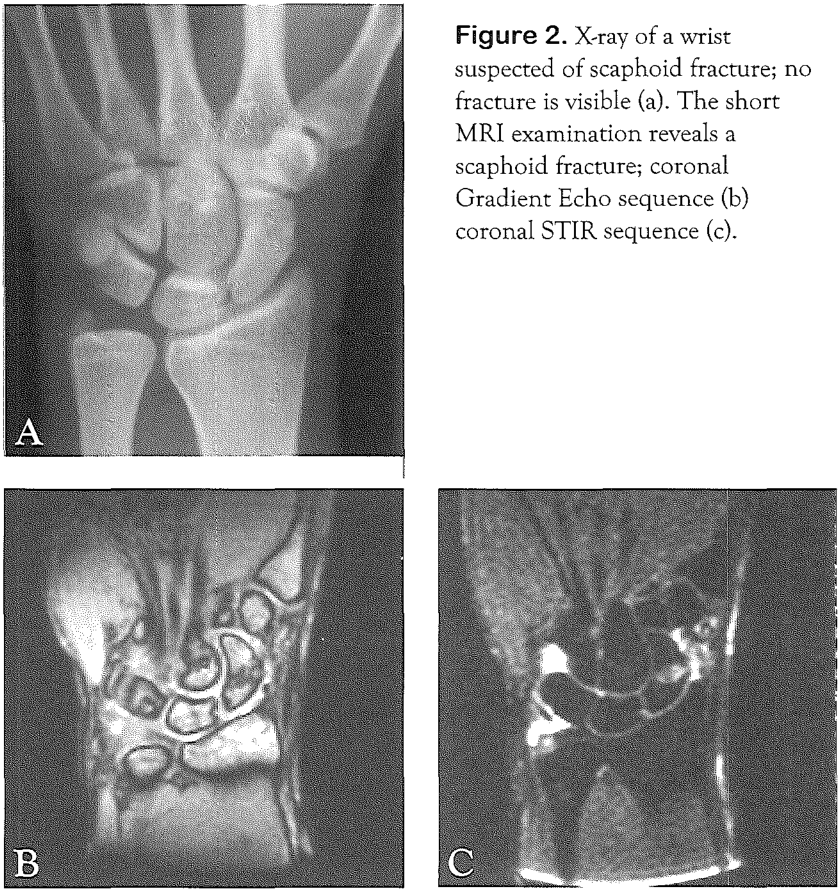


Figure 2. X-ray of a wrist suspected of scaphoid fracture; no fracture is visible (a). The short MRI examination reveals a scaphoid fracture; coronal Gradient Echo sequence (b) coronal STIR sequence (c).

overmodelling. Tenderness of the anatomical snuffbox, radiography result and MRI result were used in the multivariable logistic regression analysis (Table 4). The Hosmer-Lemeshow goodness of fit test showed a non-significant p-value for all the models, indicating good calibration of the models.

In all the models the snuffbox test demonstrated a relative odds ratio of more than 2 in predicting the outcome, but statistical significance could not be demonstrated (likelihood ratio test p-value >0.05, stepwise backward approach). Radiography, MRI and the combined 'imaging variable' contributed significantly in predicting the outcome in all the models. In the model with radiography and MRI (model 3) a negative MRI result did not predict that treatment was unnecessary (since the confidence interval of the odds ratio contained 1), whereas a positive result significantly predicted the need for treatment.

Also in the combined 'imaging variable' the combination of radiography and MRI both negative had no significant predictive value. All other combinations in the combined 'imaging variable' were significantly predictive of the need for treatment.

TABLE 3. Predictors of wrist treatment - univariable regression analysis

Variable (predictor)	Proportion treated in group with predictor	Proportion treated in reference group	Relative Odds ratio (95% CI)	chi-square p-value¶
Sex (male)	18/52 (35%)	18/35 (51%)	0.50 (0.21-1.20)	0.12
Age (continuous)			1.02 (1.00-1.05)	0.07
Side (right vs left)	18/43 (42%)	18/44 (41%)	1.04 (0.44-2.44)	0.93
Side (right vs left)	18/43 (42%)	18/44 (41%)	1.04 (0.44-2.44)	0.93
Snuffbox tenderness present (vs absent or unknown)	14/25 (56%)	22/62 (35%)	2.31 (0.90-5.96)	0.08
X-ray pos (vs neg or uncertain*)	26/30 (87%)	10/57 (18%)	31.2 (8.90-109)	<0.01
MRI result (vs not made)				0.06
MRI neg	5/24 (21%)	21/47 (45%)	0.39 (0.13-1.16)	
MRI pos or uncertain**	10/16 (63%)	21/47 (45%)	1.86 (0.57-6.06)	
Combined 'imaging variable' †				<0.01
X-ray and MRI both negative	3/20 (15%)	1/23 (4%)	0.35 (0.33-37.0)	
X-ray positive, MRI not made	18/24 (75%)	1/23 (4%)	89.8 (9.18-878)	
X-ray and MRI discrepant ‡	8/15 (53%)	1/23 (4%)	22.8 (2.41-216)	
X-ray and MRI both positive	4/5 (80 %)	1/23 (4%)	79.8 (4.09-1558)	

¶ variables with p-value < 0.10 were included on the multivariable analyses
* 3 X-ray results were uncertain
** 3 MRI results were uncertain.
† Reference category: radiography negative and MRI not made
‡ Discrepancy between radiography and MRI result: 4 times radiography positive and MRI negative, 11 times MRI positive and radiography negative.

TABLE 4. Predictors of wrist treatment in multivariable analysis - relative odds ratios and model performance statistics.

	Model 1 Snuffbox test and radiography (95 % CI)	Model 2 Snuffbox test and MRI (95 % CI)	Model 3 Snuffbox test, radiography and MRI (95 % CI)	Model 4 Snuffbox test and combined 'imaging variable' (95 % CI)
Snuffbox tenderness				
present (vs absent or unknown)	2.65 (0.81-8.64)	2.16 (0.80-5.84)	2.29 (0.65-8.08)	2.11 (0.57-7.78)
X-ray result positive				
(vs uncertain or negative)	17.1 (5.64-52.2)		24.7 (6.59-93.1)	
MRI result (vs no MRI made)				
negative		0.32 (0.10-1.01)	0.87 (0.20-3.82)	
positive or uncertain		1.83 (0.56-6.02)	6.28 (1.27-31.0)	
Combined 'imaging variable'*				
X-ray and MRI both negative				3.53 (0.33-37.6)
X-ray positive, MRI not made				105 (10.8-1040)
X-ray and MRI discrepant				24.6 (2.58-235)
X-ray and MRI both positive				69.5 (3.45-1400)
Model performance statistics				
Hosmer-Lemeshow GOF test	p=0.66	p=0.86	p=0.47	p=0.71
Akaike's Information Criterion	30.8	3.97	33.7	36.8
Area under ROC curve	0.83	0.70	0.87	0.88
(95% CI)	(0.74-0.92)	(0.58-0.81)	(0.78-0.95)	(0.81-0.95)

CI = confidence interval, GOF = goodness of fit.

* Reference category: radiography negative and MRI not made

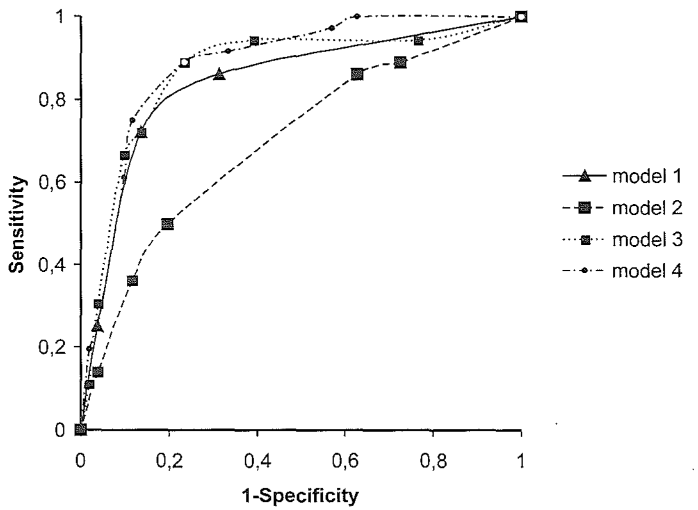


Figure 3. Receiver Operating Characteristic Curves of the multivariable models:
Model 1 : snuffbox tenderness and radiography
Model 2 : snuffbox tenderness and MRI
Model 3 : snuffbox tenderness, radiography and MRI
Model 4 : snuffbox tenderness and the combined 'imaging variable' (containing radiography and MRI)

AIC of the model with both radiography and MRI (model 3) was significantly larger than AIC of the models with only radiography (model 1) or only MRI (model 2)(likelihood ratio test p-value 0.03 and <0.001 respectively). The difference in AIC between model 3 and model 4 could not be assessed statistically, since these models are not nested, but since the difference was more than 2 it was considered significant. The areas under the ROC curve - representing discriminatory power - of all the models containing radiography did not differ significantly (table 4 and figure 3), however, they were all significantly larger than the area under the ROC curve of the model without radiography (model 2).

Discussion

In search for improvement of the early diagnostic accuracy in patients presenting with acute wrist trauma, we assessed the predictive value of radiography and MRI in predicting the need for additional treatment. The data for this study were obtained from an RCT assessing the effectiveness of a short MRI examination in all patients with acute wrist trauma and used to create four prediction models.

Although one doesn't need to perform an RCT to create prediction rules, the data can be used for this purpose.
The addition of MRI to radiography resulted into a significant increase in predictive power in the identification of patients that would need additional treatment. The increase of predictive power was only attributable to a positive MRI result, since a negative MRI result did not add to the information obtained from the radiograph in predicting that treatment was unnecessary. This means that if the radiograph is negative, a negative MRI result does not add significant information, whereas a positive MRI result significantly improves the predictive value. For the model with radiography and MRI as a whole, the discriminatory power did not differ significantly from the model with only radiography. MRI instead of radiography decreased the predictive power.
The only additional contribution of MRI in our population was in identifying radiographically occult fractures of the scaphoid bone and the distal radius. All of the scaphoid fractures and one of the distal radius fractures were treated with prolonged cast. The immobilization of the distal radius fracture was prolonged because of persisting pain. In contrast to scaphoid fractures, occult distal radius fractures are not known for potential complications if they are not diagnosed. Our results suggest that potential value of initial MRI in the acutely injured wrist is only to be expected in clinically suspected but radiographically occult scaphoid fractures. The number of scaphoid fractures in our study was to small to make a definitive statement regarding the value of MRI to exclude scaphoid fracture in the face of snuffbox tenderness. The value of MRI for the detection of scaphoid fractures has been shown by others (4, 8, 18-19). Some studies indicate that early MRI in these patients can be cost-effective from a hospital perspective (4, 19) and potentially also from a societal perspective (18).
The addition of a short dedicated extremity MRI examination to the initial diagnostic strategy did not lead to the diagnosis of soft tissue injury requiring treatment. The TFCC lesions and scapholunate ligament rupture as diagnosed with MRI did not lead to treatment or persistent symptoms.

Limitations of this study

In our study population no patient was treated because of soft tissue injury, therefore we could not demonstrate any predictive value of initial MRI in this respect. Our results suggest that the incidence of soft tissue injury in wrist trauma is relatively low, however, the observed low incidence of soft tissue injury in our study may be influenced by the relatively short follow-up period of 6 months. Especially ligamentous instability can potentially take months to years to be recognized

clinically. A low incidence of soft tissue injury after wrist trauma was also found by Raby in a study on MRI of suspected scaphoid fractures, (19). It must be noted that many studies report a low sensitivity of MRI for the detection of scapholunate ligament rupture, although the specificity is generally high (9-12).
Another limitation of this study is the fact that a considerable number of eligible patients were inadvertently missed. The main reason for missing patients was that inclusion took place from 8.00 - 23.00 hour, seven days a week. Consequently a large number of radiology technologists were involved in recruitment and not all were equally aware of the importance of consecutive recruitment and many were forgetful.
We assumed that the clinical decision whether or not to treat a patient was always correct. However, it is possible that some patients were treated unnecessarily or that patients were inadvertently not treated. The first group is difficult to estimate; regarding the latter group, we found that four patients without a demonstrated lesion on radiography or MRI still had complaints of the injured wrist three months after the injury. One of these patients was strongly suspected of having a borderline personality disorder and had visited our hospital for numerous complaints. Another patient received physical therapy, without a clear diagnosis. The other two patients received physical therapy, but did not seek medical attention for their persistent symptoms. It may be possible that these last three patients in fact had a significant, but undetected lesion. This may be caused by the fact that we used an MRI protocol with a limited matrix size and limited number of excitations to reduce scanning time. However, if we missed a lesion in these patients, it apparently was no lesion that led to persisting complaints without treatment, since all three were symptomfree within 6 months.

The strategy of a short MRI examination in the initial evaluation of acute wrist trauma may be implemented in daily routine only if it has shown to be cost-effective. The potential is to reduce unnecessary treatment and to expedite necessary treatment, thereby reducing time off work, which is an important cost to society. A reduction of these costs may outweigh the costs of a short MRI examination. Further work should focus on this issue.

Conclusion

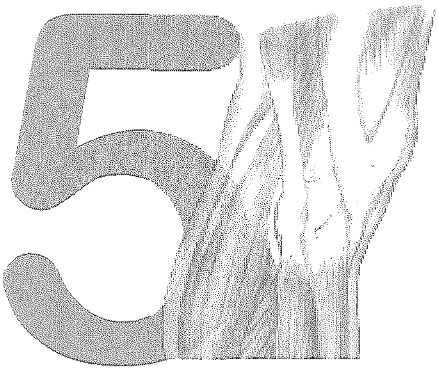
A short MRI examination on a low field MRI system in the initial workup of all patients with acute wrist trauma has additional value in predicting the need for treatment in addition to radiography, but it does not add in the identification of patients who can be safely discharged from follow-up.

Early MRI may, however, be valuable when performed in patients clinically suspected of a scaphoid fracture, but with normal radiographs.

References

1. Breitenseher MJ, Trattnig S, Gabler C, Happel B, Bankier A, Kukla C, et al. [MRI in radiologically occult scaphoid fractures. Initial experiences with 1.0 Tesla (whole body-middle field equipment) versus 0.2 Tesla (dedicated low-field equipment)]MRT bei radiologisch okkulten Kahnbeinfrakturen. Erste Erfahrungen von 1.0 Tesla (Ganzkorper-Mittelfeldgerat) versus 0.2 Tesla (dediziertes Niederfeldgerat). Radiologe 1997;37(10):812-8.
2. Tiel-van Buul MM, van Beek EJ, Broekhuizen AH, Nooitgedacht EA, Davids PH, Bakker AJ. Diagnosing scaphoid fractures: radiographs cannot be used as a gold standard! Injury 1992;23(2):77-9.
3. Langhoff O, Andersen JL. Consequences of late immobilization of scaphoid fractures. J Hand Surg [Br] 1988;13(1):77-9.
4. Dorsay TA, Major NM, Helms CA. Cost-effectiveness of immediate MR imaging versus traditional follow-up for revealing radiographically occult scaphoid fractures. AJR Am J Roentgenol 2001;177(6):1257-63.
5. Breitenseher MJ, Metz VM, Gilula LA, Gaebler C, Kukla C, Fleischmann D, et al. Radiographically occult scaphoid fractures: value of MR imaging in detection. Radiology 1997;203(1):245-50.
6. Hunter JC, Escobedo EM, Wilson AJ, Hanel DP, Zink-Brody GC, Mann FA. MR imaging of clinically suspected scaphoid fractures. AJR Am J Roentgenol 1997;168(5):1287-93.
7. Kitsis C, Taylor M, Chandey J, Smith R, Latham J, Turner S, et al. Imaging the problem scaphoid. Injury 1998;29(7):515-20.
8. Bretlau T, Christensen OM, Edstrom P, Thomsen HS, Lausten GS. Diagnosis of scaphoid fracture and dedicated extremity MRI. Acta Orthop Scand 1999;70(5):504-8.
9. Kato H, Nakamura R, Shionoya K, Makino N, Imaeda T. Does high-resolution MR imaging have better accuracy than standard MR imaging for evaluation of the triangular fibrocartilage complex? J Hand Surg [Br] 2000;25(5):487-91.
10. Schadel-Hopfner M, Iwinska-Zelder J, Braus T, Bohringer G, Klose KJ, Gotzen L. MRI versus arthroscopy in the diagnosis of scapholunate ligament injury. J Hand Surg [Br] 2001;26(1):17-21.
11. Hobby JL, Tom BD, Bearcroft PW, Dixon AK. Magnetic resonance imaging of the wrist: diagnostic performance statistics. Clin Radiol 2001;56(1):50-7.
12. Johnstone DJ, Thorogood S, Smith WH, Scott TD. A comparison of magnetic resonance imaging and arthroscopy in the investigation of chronic wrist pain. J Hand Surg [Br] 1997;22(6):714-8.
13. Hosmer D, Lemeshow S. Applied logistic regression. New York: John Wiley & Sons, Inc.; 1989.

14. Hunink MGM, Glasziou PP, Siegel JE, al e. Decision making in health and medicine: integrating evidence and values. Cambridge, England: Cambridge University Press; 2001.
15. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 1983;148(3):839-43.
16. Akaike H, editor. Information theory as an extension of the maximum likelihood principle. Budapest: Akademia Kiado; 1973.
17. Scheck RJ, Kubitzek C, Hierner R, Szeimies U, Pfluger T, Wilhelm K, et al. The scapholunate interosseous ligament in MR arthrography of the wrist: correlation with non-enhanced MRI and wrist arthroscopy. Skeletal Radiol 1997;26(5):263-71.
18. Gaebler C, Kukla C, Breitenseher M, Trattnig S, Mittlboeck M, Vecsei V. Magnetic resonance imaging of occult scaphoid fractures. J Trauma 1996;41(1):73-6.
19. Raby N. Magnetic resonance imaging of suspected scaphoid fractures using a low field dedicated extremity MR system. Clin Radiol 2001;56(4):316-20.
20. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 1996;49(12):1373-9.



Traumatic Knee Injury: the Value of A Short Dedicated Extremity MRI Examination in the Early Stage in Predicting Subsequent Treatment

Edwin H.G. Oei, MSc [1,2,3], Jeroen J. Nikken, MD [1,2,3], Abida Z. Ginai, MD, PhD [2], Gabriel P. Krestin, MD, PhD [2], Jan A.N. Verhaar, MD, PhD [4], Arie B. van Vugt, MD, PhD [5], M.G. Myriam Hunink, MD, PhD [1,2,3]

From the Program for the Assessment of Radiological Technology (ART Program)[1], Department of Radiology [2] and Department of Epidemiology and Biostatistics [3]; Department of Orthopaedic Surgery [4] and Department of Traumatology [5].
Erasmus MC, University Medical Center Rotterdam.

Submitted for publication.

Abstract

Purpose
To assess the predictive value of dedicated extremity MRI in the initial stage after knee trauma in distinguishing patients requiring treatment following the first hospital visit from those who can be discharged from follow-up.

Materials and Methods
196 patients presenting with recent knee injury were included in a prospective randomized clinical trial if a plain radiograph was ordered. They were randomized between radiography only and radiography plus immediate MRI as diagnostic strategy. MRI was performed on a 0.2 Tesla dedicated extremity MRI scanner using four short pulse sequences. Univariable and multivariable logistic regression analysis was used to evaluate patient characteristics, trauma mechanism, and findings on radiography and MRI for their ability to predict whether treatment was required within the follow-up period of six months after initial presentation.

Results
Treatment was required in 109 patients. Age 30 years or higher, indirect trauma mechanism, radiography result, and MRI result were significant predictors of treatment in both the univariable and multivariable analysis ($p<0.05$). However, in the multivariable analysis a normal MRI result did not contribute significantly to the prediction that additional treatment would be unnecessary. Only an abnormal MRI result significantly predicted the need for treatment and only when MRI was used as a replacement of radiography (odds ratio 2.61 [95% CI: 1.12-6.06]).

Conlusion
Implementation of a dedicated extremity MRI examination in addition to or instead of plain radiography improves the prediction whether additional treatment is required in patients presenting with traumatic knee injury, but does not significantly aid in identifying patients who can be discharged from further follow-up. Therefore, the additional value of a short MRI examination in the initial stage after knee trauma is limited.

Introduction

Traumatic injuries of the knee joint, often caused by sports activities, constitute a large proportion of musculoskeletal trauma encountered in the emergency department. The initial evaluation of knee injuries usually consists of taking the clinical history and performing a physical examination, involving various manipulative tests (1, 2). Although physical examination may aid in establishing the diagnosis at a later stage or when performed under anaesthesia, its accuracy in the acute stage has been questioned especially for meniscal tears, and it is influenced by many factors (3-5). Moreover, it has been well recognized that thorough physical examination of a recently injured knee with acute hemarthrosis is often difficult due to swelling, pain, and guarding (6-8).

If a fracture is suspected, plain radiography is usually added to the diagnostic work-up, often based on criteria previously published as the Ottawa Knee Rule (9, 10). However, traumatic abnormalities other than a fracture may be present in the knee joint that cannot be visualized by radiography. Unlike radiography, magnetic resonance imaging (MRI) does provide information on soft tissue damage and has widely been accepted for the evaluation of internal knee derangements. A recent systematic review and meta-analysis demonstrated that MRI has a very good performance in diagnosing tears of the menisci and cruciate ligaments (11). Nonetheless, the routine use of MRI in the initial work-up of acute knee injuries, has been hampered by the high costs, the limited availability, and the long duration of an MRI examination. Low-field dedicated extremity MRI scanners have been developed to overcome some of these problems. Because of the compact design and low-field strength of these type of scanners, the costs can be kept relatively low compared to high-field whole-body units. Therefore, low-field dedicated extremity MRI has created the possibility to apply MRI more routinely in the initial evaluation of knee trauma.

Application of MRI as an initial examination tool after knee trauma in the emergency room setting can potentially yield benefit to the patient and the health care system. Detailed imaging information in an early stage may result in more timely diagnosis and treatment in patients who would otherwise have been followed with delay of definitive therapy. Conversely, patients could possibly be identified who are unlikely to require specific treatment and who will thus not need reassessment or follow-up. In the light of cost minimization and efficiency this may be beneficial since these patients do not need to return to the outpatient department.

Purpose of this study was to assess the predictive value of MRI in the initial stage after knee trauma, additionally to and instead of radiography, in distinguishing patients requiring treatment from those who can safely be discharged without additional work-up or follow-up. For this purpose we implemented a low-field dedicated extremity MRI examination using a short protocol.

Materials and methods

Study design and study population

This study was part of a prospective RCT aimed at assessing the value and cost-effectiveness of dedicated extremity MRI in the initial evaluation of all patients with acute knee injury. Although it is not necessary to perform an RCT to study the predictive value of MRI in acute knee trauma, we used the data from such a trial to do so. Where applicable we followed published guidelines for reporting the results of RCTs (12) and otherwise those for reporting diagnostic studies (13). The study was approved by our institutional review board.

From August 1999 to May 2001 we recruited patients from the emergency department of our university hospital who were referred for plain radiography by a traumatologist or orthopaedic surgeon because of traumatic knee injury within the preceding seven days. We excluded patients who had additional injury of the head, back, thorax, or abdomen, patients with a compound knee fracture, patients requiring urgent treatment (e.g. in case of threatened circulation of the leg), and patients with pre-existing symptoms of the same knee. Patients were included seven days per week from 08.00 a.m. to 11.00 p.m. All eligible patients were provided with written and oral information about the aim of the study. After informed consent was obtained, patients were randomized between the current diagnostic work-up (only plain radiography) and a strategy in which plain radiography was immediately followed by a short dedicated extremity MRI examination. Randomization was performed by research staff not involved in the further management of the patient, using consecutively numbered sealed envelopes containing computer-generated random assignments. We applied block randomization with a block size of 20 to achieve equal numbers of patients in both study arms.

Imaging technique and interpretation

Radiography was performed in anteroposterior and lateral direction, complemented by additional patellar or tunnel views if a patellar fracture or a loose body was suspected by the examining physician. For the MRI examination, we used a low-field dedicated extremity MRI system (Artoscan M, Esaote Biomedica, Genoa, Italy) with a 0.2 Tesla permanent magnet and a dual phased array knee coil. A larger linear knee coil was used in case of a very large knee. Regarding the scanning protocol we aimed at providing a quick examination in the emergency setting, minimizing costs and burden to the patient. With these goals in mind we adapted several scanning parameters in order to reduce total examination time. This resulted in some loss of image quality, but in a pilot study the scanning protocol as listed in Table 1 was found to be of acceptable quality to detect most lesions (14).

Table 1: MRI scanning protocol

	Plane	TR (ms)	TE (ms)	Matrix	Slice thickness (mm)	Acquisition time (range; min:sec)
(Half Fourier) SE T1-w	Sagittal	490-720	18-24	192 x 216	5.0	0:49 - 1:08
TME PD-w and T2-w	Coronal	2150-2640	28-38/90	192 x 128	5.0	2:23 - 2:41
GE *	Coronal	500-595	15-16	192 x 128	5.0	0:48 - 0:53
STIR†	Sagittal	1060-1290	24-28	192 x 128	6.0/7.0	1:37 - 1:46

Note: The field of view was 200x140 mm and the number of excitations was 1 for all sequences.
SE = spin echo, TME = turbo multi echo, PD = proton density, GE = gradient echo, STIR = short-tau inversion recovery, TR = repetition time, TE = echo time.
* Flip angle = 75 degrees,
† Inversion time = 80-85 ms

Average acquisition time was 6:07 minutes with a total examination time, including start-up of the MRI unit and patient positioning, of less than 15 minutes. According to daily practice in our hospital, the radiographs were initially assessed by the treating physician in the emergency department and they were later reassessed by a musculoskeletal radiologist during a daily plenary session with the treating physician. The MRI was immediately interpreted by an experienced musculoskeletal radiologist (A.Z.G.) who was unaware of the radiography result.

The findings on MRI were reported to the treating physician in the emergency department, so that they could aid in the decision whether to treat or follow the patient. In a minority of cases (e.g. during the weekends and in the evenings) the MRI was assessed by a radiology resident and reassessed by the musculoskeletal radiologist the next day. The treating physician was notified of any discrepancies in interpretation.

For the purpose of our analysis, the radiography and MRI results were scored as abnormal or normal, based on the presence or absence of recent traumatic abnormalities that generally would require treatment. Radiography was considered abnormal if a fracture, epiphysiolysis, dislocation, or osteochondral lesion was visible. The MRI result was scored as abnormal if one or more of the following lesions were present: fracture, epiphysiolysis, dislocation, osteochondral lesion, meniscal tear, anterior cruciate ligament (ACL) total rupture, posterior cruciate ligament (PCL) total rupture, medial collateral ligament (MCL) total or partial rupture, and lateral collateral ligament (LCL) total or partial rupture. Although a partial rupture of the ACL or PCL is obviously a traumatic lesion, we considered the MRI result as normal for the purpose of our analysis if this was the only visible abnormality, since an isolated partial tear of a cruciate ligament does not usually need specific treatment. A partial collateral ligament rupture, however, is usually treated with a knee brace. Similarly, radiography and MRI were scored as normal (i.e. not requiring specific treatment) if only joint effusion, hemarthrosis, soft tissue edema, bone marrow edema, meniscal degeneration or osteoarthritis was seen, or if the findings suggested an old traumatic lesion.

Outcome and follow-up

The outcome measure was additional treatment within six months after the initial visit to the emergency department. Treatment was defined as any specific additional therapy, including surgery, physiotherapy with supervised and structured rehabilitation exercises, plaster cast, and taping. Minor initial treatment measures, such as a pressure bandage, were not regarded as a specific treatment.

Patients were followed as long as they had daily complaints of the affected knee up to a maximum of six months after initial presentation. Data were collected by searching the emergency department records, hospital records, outpatient clinic records, and the electronic hospital information system. In addition, we sent questionnaires to all patients, asking whether they had undergone treatment in another institution.

These questionnaires were mailed at one week, six weeks, twelve weeks, and six months after the initial visit to the emergency department. If the questionnaires were not returned we attempted to interview the patients by telephone. The final diagnosis was determined by reviewing all information from diagnostic imaging, follow-up, questionnaires, or arthroscopy or operation reports if available.

Data analysis

The data were analyzed by intention-to-diagnose-and-treat using univariable and multivariable logistic regression analysis. We evaluated the following independent variables for their ability to predict whether treatment was required within 6 months after the initial hospital visit: age (both continuous and dichotomized at various threshold levels), sex, side, trauma mechanism (indirect or direct trauma), and the result of both radiography and MRI (abnormality requiring treatment versus no abnormality). The absence of MRI information, which was the case in half of our patient population due to the randomized study design, was used as the reference category for the MRI result in the analysis (normal or abnormal versus no MRI information). This allowed us to determine the additional value of the MRI result compared to no MRI information.

To account for the fact that the findings on radiography and MRI are correlated, we created an overall 'imaging variable' integrating the results of both the radiography and MRI examination. In this variable, the following five possible combinations of radiography and MRI results were coded using four dummy variables: 1) radiography normal and no MRI information (reference category), 2) both radiography and MRI normal, 3) radiography abnormal and no MRI information, 4) radiography and MRI discrepant, and 5) radiography and MRI both abnormal. Using this 'imaging variable' one takes into account the diagnostic interaction between radiography and MRI, which may be involved if two imaging modalities are used together to visualize the same lesion. It is therefore a more appropriate manner of analyzing the result of radiography and MRI combined, rather than using the two variables separately in the same regression model.

We first assessed each variable separately using univariable logistic regression analysis and calculated the odds ratio (OR) and the 95% confidence interval (CI). Subsequently, we performed multivariable logistic regression analyses including all variables that had a chi-square p-value of less than 0.10 in the univariable analysis. Because we wanted to assess the value of the imaging modalities we created four models with inclusion of respectively radiography (Model 1), MRI (Model 2), radiography and MRI (Model 3), and the overall 'imaging variable' described above (Model 4). We used Hosmer and Lemeshow's goodness-of-fit test to evaluate the calibration of the models. A Hosmer and Lemeshow goodness-of-fit test that is not significant (p-value > 0.05) indicates that the model fits well with the observed data (15). The area under the receiver operating characteristic (ROC) curve was calculated to assess the discriminatory power of each model in distinguishing patients requiring treatment from those who can be discharged. A larger area under the ROC curve indicates higher discriminatory power (16). To compare the area under the ROC curve across the models, we used the method described by Hanley and McNeil (17), which adjusts for the fact that the models are correlated since they were derived from the same sample of patients. The likelihood ratio test, which can be used for comparing nested models, was used to assess the value of MRI in addition to radiography by comparing Model 1 and Model 3. To compare models that are not nested, one needs to calculate Akaike's infomation criterion (AIC) for each model and assess the difference in AIC between the models. The AIC of a model is calculated by subtracting two times the degrees of freedom from the model chi-square estimate (18). A higher AIC indicates better model performance. We used this technique to assess the value of MRI as a replacement of radiography by comparing the AIC of Model 1 with Model 2, and also in addition to radiography by comparing the AIC of Model 1 with Model 4. A difference in AIC larger than 2 was considered significant. All analyses were performed using SPSS for Windows (release 10.0.0; SPSS Inc., Chicago, IL).

Sample size

The data for this study were obtained from an RCT evaluating the costs and effectiveness of early MRI of knee trauma and, hence, the number of independent variables that could be analyzed in the multivariable regression analysis was determined by the number of patients included in the RCT. 109 included patients received treatment and 80 patients did not need receive treatment. Since at least 10 patients with an event and 10 patients without an event are necessary for each independent variable in the multivariable logistic regression analysis (19), a maximum of 8 independent variables could be analyzed.

Results

196 patients were initially included in the study. Although inclusion was intended to be consecutive, only about half of the eligible patients were included, the main reason being that radiology technologists on service who were supposed to identify eligible patients and ask them to participate in the study often forgot to do so.

Six patients were incorrectly included and they did not fulfill the inclusion criteria on reascertainment. One patient presented to our hospital more than seven days after knee injury, three had a history of pre-existing symptoms of the knee joint, and another two patients had a non-traumatic origin of knee pain. These six patients, together with one patient in whom automutilation was strongly suspected, were excluded from the analysis. Of the 189 remaining patients (123 male, 66 female; mean age 33.4 years, standard deviation 13.9, range 12.6-74.6) 96 patients were randomized to radiography as the diagnostic strategy, while in 93 patients radiography was followed by MRI. A flow diagram of subjects passing through the study is presented in Figure 1.

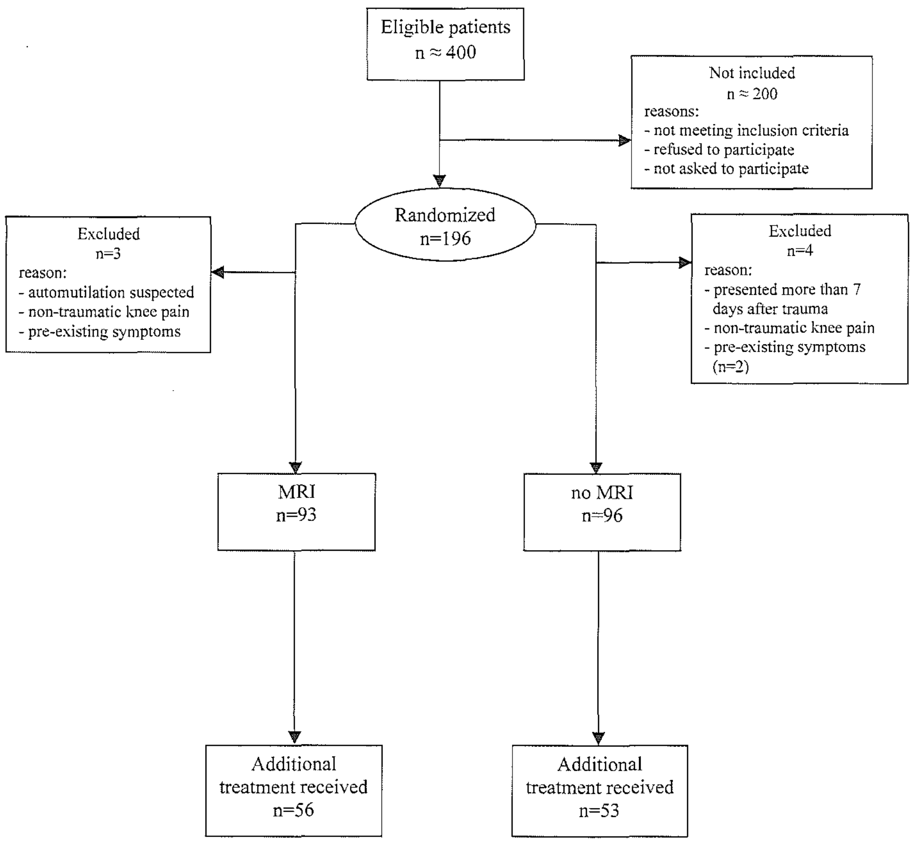


Figure 1. Flow diagram of patients passing through the study.

The lesions that were found on MRI together with the corresponding frequencies are listed in Table 2 (see Figures 2 and 3 for illustrative examples).

Table 2: Frequency of lesions found on MRI

Lesion type	Frequency
Fracture	11
Anterior cruciate ligament total rupture	17
Anterior cruciate ligament partial rupture *	9
Posterior cruciate ligament total rupture	0
Posterior cruciate ligament partial rupture *	1
Medial collateral ligament total rupture	2
Medial collateral ligament partial rupture	8
Lateral collateral ligament total rupture	1
Lateral collateral ligament partial rupture	5
Medial meniscus rupture	17
Lateral meniscus rupture	10
Osteochondral lesion	3
Bone marrow edema *	23

Note: multiple lesions may be present in the same patient
* not regarded as a traumatic lesion requiring treatment in the analysis

During the follow-up period 109 patients underwent treatment, consisting of arthroscopic surgery in 21 cases including ACL reconstruction in 1 case, open surgery in 5 cases, drainage in 1 case, and plaster cast immobilization in 23 cases, sometimes supplemented by physiotherapy. Among the 21 cases of arthroscopic surgery, partial meniscectomy was performed in 18 cases. Of these patients, 10 patients had been randomized to the MRI diagnostic strategy and in all these patients the meniscal tear was diagnosed on MRI. The 5 cases of open surgery consisted of resection of an ACL fragment combined with reefing of the MCL, rectus femoris tendon suturing, osteosynthesis of a tibial plateau fracture and spongiosa-plasty, reinsertion of the MCL on the tibia, and fixation of an osteochondritis dissecans fragment with bone pins. In 56 patients physiotherapy was the only treatment.

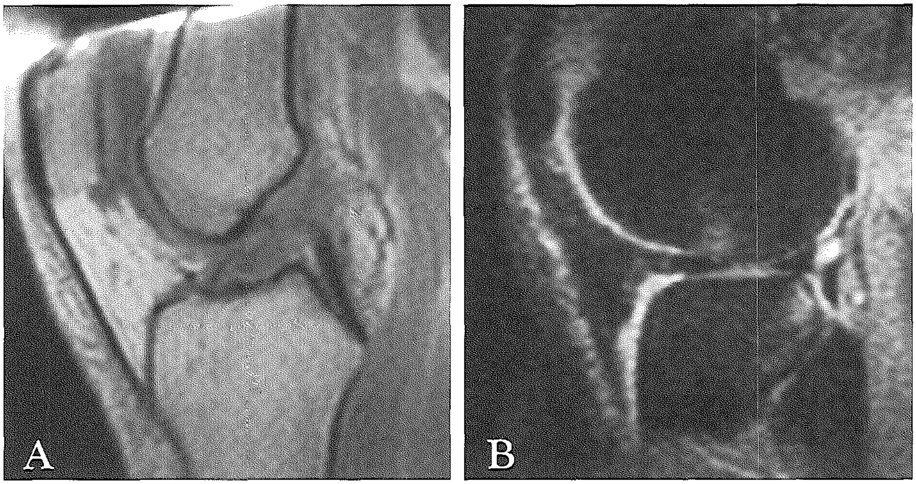


Figure 2a/b.
a) Anterior cruciate ligament tear (sagittal T1-weighted Spin Echo sequence).
b) On the STIR sequence associated bone marrow edema is visible in the lateral femoral condyle and the posterolateral tibial plateau.

The ORs and 95% CIs of all demographic, clinical, and imaging variables that were assessed in the univariable logistic regression analysis are presented in Table 3. Age, trauma mechanism, radiography result, MRI result, and the overall 'imaging variable' had a chi-square p-value of less than 0.10 and were subsequently used in the multivariable analysis. Age of the patient was a significant predictor of treatment as continuous variable, but the most significant association was found when age was dichotomized with 30 years as threshold. Compared to no MRI information, an abnormal MRI result was predictive for the need for treatment, whereas a normal MRI result was predictive of no treatment required. Similarly, regarding the overall 'imaging variable', the category in which radiography and MRI were both abnormal was significantly predictive for treatment, and the category in which radiography and MRI were both normal was significantly predictive of no treatment required. The two intermediate categories had no significant association with treatment. As expected, sex and side were insignificant in the univariable analysis.

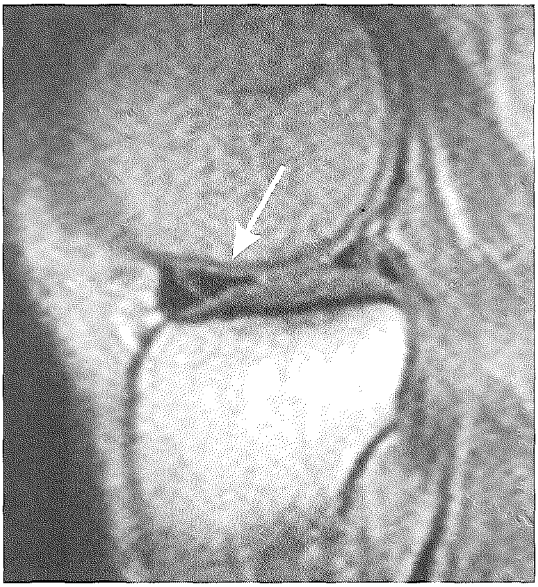


Figure 3. Tear of the posterior horn of the lateral meniscus (sagittal T1-weighted Spin Echo sequence). A displaced fragment is seen posterior to the anterior horn (arrow).

We created four multivariable models including the significant variables from the univariable analysis, with different combinations of the two imaging modalities (Table 4). All variables were significantly contributive to these models (all likelikood ratio test p-values < 0.05). Model 1 consisted of age, trauma mechanism, and the result of radiography. In Model 2 the result of radiography was replaced by the MRI result, which was only found to be a significant predictor of the need for treatment if abnormal, whilst a normal MRI result could not predict that treatment was not required, reflected by a 95% CI of the odds ratio containing unity. Model 3 consisted of age, trauma mechanism, and the results of both imaging modalities using the separate variables for radiography and MRI results (additive combination of Models 1 and 2). In this model the odds ratios of the MRI variable were insignificant for both a normal and an abnormal MRI result, suggesting that a positive or a negative MRI result, versus no MRI information available, does not significantly contribute to the prediction of the need for treatment when controlled for age, trauma mechanism, and radiography result. Model 4 consisted of age, trauma mechanism, and the overall 'imaging variable'. Although this variable did significantly contribute in the prediction of treatment (p=0.002), only the odds ratio of the category in which both radiography and MRI were abnormal was significant.

Table 3: Performance of clinical and imaging variables in predicting the need for treatment - univariable logistic regression analysis

Variable	Proportion treated in group with predictor	Proportion treated in reference group	Univariable Odds ratio (95% CI)	chi-square p-value*
Age (continuous)			1.02 (1.00-1.04)	0.06
Age ≥ 30 years (versus < 30 years)	64/95 (67%)	45/94 (48%)	2.25 (1.25-4.05)	0.01
Sex (male versus female)	68/123 (55%)	41/66 (62%)	0.75 (0.41-1.39)	0.36
Side (left versus right)	49/88 (56%)	60/101 (59%)	0.86 (0.48-1.53)	0.61
Indirect trauma (versus direct trauma) †	76/109 (70%)	32/78 (41%)	3.31 (1.80-6.09)	<0.001
Radiography: abnormal (versus normal) ‡	19/24 (79%)	88/161 (55%)	3.15 (1.12-8.85)	0.02
MRI result (no MRI information as reference)				<0.001
normal	15/42 (36%)	53/96 (55%)	0.45 (0.21-0.95)	
abnormal	41/51 (80%)	53/96 (55%)	3.33 (1.49-7.40)	
Imaging variable § ‡				<0.001
radiography and MRI both normal	14/40 (35%)	47/85 (55%)	0.44 (0.20-0.95)	
radiography abnormal, no MRI info	5/9 (56%)	47/85 (55%)	1.01 (0.25-4.03)	
radiography and MRI discrepant	28/37 (76%)	47/85 (55%)	2.52 (1.06-5.97)	
radiography and MRI both abnormal	13/14 (93%)	47/85 (55%)	10.5 (1.32-84.0)	

CI = confidence interval: the odds ratio can be considered significant if the 95% CI does not include unity.

* Variables with p-value < 0.10 were included in the multivariable analysis

† Analysis of 187 patients since data on 2 patients were missing

‡ Analysis of 185 patients since data on 4 patients were missing

§ Reference category: Radiography negative and MRI not made

Table 4: Multivariable prediction models for treatment

	Odds ratio (95% CI)	Likelihood ratio test p-value		
Predictor	Model 1	Model 2	Model 3	Model 4
Age (≥ 30 years versus < 30 years)	2.88 (1.47-5.62) p=0.001	2.38 (1.22-4.63) p=0.010	2.62 (1.32-5.21) p=0.005	2.52 (1.26-5.03) p=0.008
Trauma (indirect versus direct)	4.89 (2.45-9.76) p<0.001	3.40 (1.74-6.63) p<0.001	4.17 (2.05-8.50) p<0.001	4.15 (2.04-8.43) p<0.001
Radiography (abnormal versus normal)	5.24 (1.67-16.4) p=0.002		3.65 (1.12-11.9) p=0.023	
MRI result (versus no MRI information)		p=0.004		p=0.041
normal		0.53 (0.24-1.19)	0.58 (0.26-1.33)	
abnormal		2.61 (1.12-6.06)	2.11 (0.89-5.00)	
Imaging variable *				p=0.002
radiography and MRI both normal				0.52 (0.22-1.21)
radiography abnormal, no MRI info				1.90 (0.43-8.39)
radiography and MRI discrepant				1.86 (0.75-4.65)
radiography and MRI both abnormal				14.0 (1.62-121)
Model performance statistics				
Hosmer-Lemeshow statistic †	1.80; p= 0.77	3.54; p=0.83	3.45; p=0.84	2.82; p=0.95
Area under ROC curve (95% CI)	0.74 (0.67-0.81)	0.74 (0.67-0.82)	0.77 (0.70-0.84)	0.77 (0.70-0.84)
Akaike's information criterion	28.44	27.71	30.84	30.07

All models fitted for the same 183 patients with no missing data.
Variables included in the models were:
Model 1: age, trauma mechanism, and radiography.
Model 2: age, trauma mechanism, and MRI.
Model 3: age, trauma mechanism, radiography, and MRI.
Model 4: age, trauma mechanism, radiography, and the 'imaging variable'
The likelihood ratio p-value concerns the predictor as a whole, whereas the odds ratios concern the predictor results.
CI = confidence interval; the odds ratio can be considered significant if the 95% CI does not include unity.
* Reference category: Radiography normal and no MRI information
† Hosmer and Lemeshow goodness-of-fit test was not significant for all models, indicating that they fit well.

All model performance statistics are tabulated in Table 4. The Hosmer and Lemeshow goodness-of-fit test was not significant for any of the models, indicating that all models fit well with the observed data. The area under the ROC curve was not significantly different across the models, suggesting that all models were equal in terms of discriminatory power.

Adding MRI to Model 1 (and thus creating Model 3) significantly improved the model performance (likelihood ratio test p-value 0.04), indicating that information from MRI improves model performance when added to radiography. However, as is noted above, this did not result in a better discrimination between patients requiring treatment and patients who can be discharged from follow-up, since both a normal and an abnormal MRI result compared to no MRI information had a 95% CI including unity. The AIC of Model 2 was found to be lower than that of Model 1, suggesting that MRI as a replacement of radiography does not improve model performance. Similarly, the performance of Model 4 (including information from MRI in the 'imaging variable') was not significantly better than that of Model 1.

Discussion

We studied the value of MRI as a routine screening tool in the initial stage after knee trauma, a setting in which MRI has not been widely implemented thus far. The purpose of performing imaging tests in the acute stage is to distinguish patients who require additional treatment from those who do not. This allows patients to be treated appropriately and others to be discharged from follow-up without the need for additional visits or work-up. MRI has been advocated as a screening tool prior to therapeutic arthroscopy, since it has proven to be very reliable in detecting internal knee derangements non-invasively (11, 20). Application of MRI as an initial examination tool in the emergency room setting, however, has been limited by the associated high costs and other problems such as burden to the patient, long scan time, and limited availability of an MRI scanner. We used a dedicated extremity MRI system, the design of which could overcome most of these problems. Because of the low-field strength and the relatively small size of the equipment it does not require infrastructural changes prior to installation, so that the costs of installation and maintenance of such an MRI scanner are low compared to a high-field whole body unit. Furthermore, it has been shown that lower magnetic field strength does not substantially reduce the diagnostic performance for most internal knee derangements (11, 21-25).

Using four multivariable logistic regression models, we assessed the value of a dedicated extremity MRI examination in predicting the need for additional treatment in patients with acute knee trauma when performed in addition to and instead of radiography. Contrary to what we expected, the results of this study suggest that such an MRI examination has limited additional value to radiography in predicting the need for treatment.

Although the MRI variable was significant in predicting the need for treatment when added to the model with radiography, compared to the current situation of no MRI information available neither an abnormal nor a normal MRI result had significant added predictive value. As an alternative to radiography, MRI was only significantly predictive of treatment if abnormal, while it did not contribute in identifying patients who will not require treatment and who can be discharged from follow-up.

We used a short MRI scanning protocol which may partly account for the limited additional value of MRI that we found in our study. However, we believe that a more expanded or routine protocol is not feasible for routine application in the emergency setting due to the excess burden to the patient because of the longer examination time and the higher costs. Furthermore, the aim was to use this protocol as a triage tool to distinguish patients who require treatment from patients who can be discharged from further follow-up. Considering this purpose, we think that a less extensive MRI technique than is commonly used for a regular MRI examination should suffice.

Our method of coding the MRI result may also have influenced the results of this study. We only considered an MRI abnormal if one or more lesions were visible that generally require therapy. Visible lesions that are usually not treated, such as a partial ACL rupture, were not counted as abnormal using this definition. Although this is justified in the majority of cases, we recognize that exceptions are possible in specific patient groups such as professional athletes. Therefore, the results of this study must be judged taking into consideration the classification method of the MRI findings. Although bone marrow edema does not require specific treatment as such, it can be an indirect manifestation of another lesion and therefore increase the likelihood of treatment. We performed a separate analysis in which isolated bone marrow edema was scored as an MRI finding requiring additional treatment, but this did not change the significance of any predictor (data not shown).
Apart from radiological techniques, we assessed the predictive value of clinical factors such as age and trauma mechanism. Higher age is associated with a higher prevalence of degenerative knee disorders, which in turn is associated with increased likelihood of structural damage after trauma. For example, a meniscal tear occurs more easily in a meniscus with myxoid degeneration than in an unaffected meniscus. In our analysis age above 30 years was found to be significantly predictive of treatment required.

Higher age is also associated with a higher prevalence of preexisting asymptomatic meniscal tears, but because of the long waiting period for arthroscopy in our institution it is unlikely that preexisting asymptomatic meniscal tears were unnecessarily treated in our study. We also recorded whether the mechanism of injury was a direct blow or an indirect, usually rotational force or varus or valgus stress, which has been shown to give rise to different patterns of findings on MRI (26, 27). Whereas a direct hit more often gives rise to a fracture, an indirect trauma more frequently causes meniscal tears, ligament ruptures, and osteochondral lesions, and was found to be a significant predictor of the need for treatment in this study.

A potential limitation of this study is that we did not assess the predictive value of findings on physical examination. However, the accuracy of physical examination of the knee joint has been controversial especially if performed in the acute stage. It has been widely recognized that extensive joint swelling, pain, or guarding render physical examination without anaesthesia inaccurate in the early stage (6-8), which was the phase of interest in this study. Moreover, it has been reported that the accuracy of physical examination of the knee is influenced by the experience of the physician (28).

Another limitation of this study is that we defined treatment within the follow-up period as the outcome measure, assuming that this truly reflected the need for treatment in all patients. However, it is possible that patients were incorrectly discharged from further treatment and follow-up, or that patients received unnecessary treatment. We believe that the assumption is justified for those patients who were analyzed as not requiring therapy. If treatment were incorrectly withheld in the initial stage, these patients would likely have returned and received treatment within the six months of follow-up because of persisting symptoms. On the other hand, the appropriateness of a treatment that has already been undergone is difficult to assess in retrospect, so that we were unable to identify patients who had been treated unnecessarily.

Follow-up of patients in this study was limited to a period of six months after their initial visit to the emergency department. Although one might argue that we may have incorrectly classified patients in whom treatment was initiated after this six months period as having undergone no treatment, we do believe that this is rarely the case, since therapy of traumatic lesions of the knee joint is generally initiated soon after the injury if necessary.

Nonetheless, we found that in some cases surgery was performed after the six months follow-up period, because of the waiting time for arthroscopy. This was the reason why we did not use surgery as an alternative outcome measure in this study. In all of these patients, however, physiotherapy had already been initiated within six months after first presentation, so that these patients were dealt with correctly in our analysis. We acknowledge that the indications for physiotherapy are often not well-defined so that it can be argued that physiotherapy should not count as a specific treatment. Therefore, we repeated the analysis in which physiotherapy as such was not considered a specific treatment. Using this definition of the outcome we found the same significant predictors in the univariable analysis. In the multivariable analysis age was only a significant predictor in Model 1, but the significance of the imaging variables in all models was not affected (data not shown). Also, all odds ratios of the imaging variables and the model performance statistics were similar, so that this additional analysis did not change the conclusions of our study.

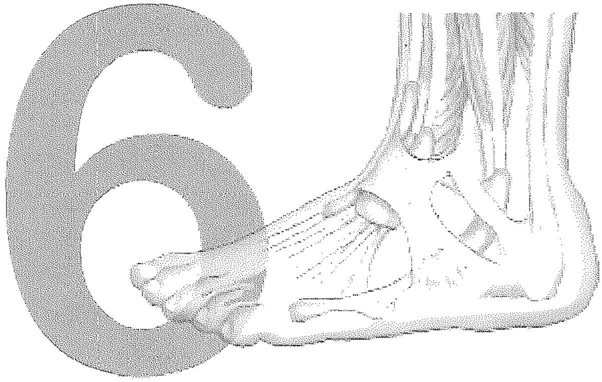
Based on the results of this study one would conclude that a prediction of the need for treatment in patients presenting after acute knee trauma can be made based on age, trauma mechanism, and the radiographic result. To apply this prediction rule in clinical practice would require choosing the optimal threshold score which requires consideration of the probability of a treatable lesion and the benefit gained by making a correct prediction weighed against the consequences of an incorrect prediction in terms of expected effectiveness and costs. The optimal threshold value of the prediction rule represents the combination of true- and false-positive ratios that yield the greatest expected utility for the patient at acceptable costs to society (16).

We conclude that the value of a dedicated extremity MRI examination in the initial stage after knee trauma is limited. Implementation of a dedicated extremity MRI examination in addition to or instead of plain radiography improves the prediction whether additional treatment is required in patients presenting with traumatic knee injury, but does not significantly aid in identifying patients who can be discharged from further follow-up.

References

1. Solomon DH, Simel DL, Bates DW, Katz JN, Schaffer JL. The rational clinical examination. Does this patient have a torn meniscus or ligament of the knee? Value of the physical examination. Jama 2001;286(13):1610-20.
2. Smith BW, Green GA. Acute knee injuries: Part I. History and physical examination. Am Fam Physician 1995;51(3):615-21.
3. Scholten RJ, Deville WL, Opstelten W, Bijl D, van der Plas CG, Bouter LM. The accuracy of physical diagnostic tests for assessing meniscal lesions of the knee: a meta-analysis. J Fam Pract 2001;50(11):938-44.
4. Kim SJ, Kim HK. Reliability of the anterior drawer test, the pivot shift test, and the Lachman test. Clin Orthop 1995(317):237-42.
5. Corea JR, Moussa M, al Othman A. McMurray's test tested. Knee Surg Sports Traumatol Arthrosc 1994;2(2):70-2.
6. Hardaker WT, Jr., Garrett WE, Jr., Bassett FH, 3rd. Evaluation of acute traumatic hemarthrosis of the knee joint. South Med J 1990;83(6):640-4.
7. Mitsou A, Vallianatos P. Clinical diagnosis of ruptures of the anterior cruciate ligament: a comparison between the Lachman test and the anterior drawer sign. Injury 1988;19(6):427-8.
8. Simonsen O, Jensen J, Mouritsen P, Lauritzen J. The accuracy of clinical examination of injury of the knee joint. Injury 1984;16(2):96-101.
9. Stiell IG, Greenberg GH, Wells GA, McKnight RD, Cwinn AA, Cacciotti T, et al. Derivation of a decision rule for the use of radiography in acute knee injuries. Ann Emerg Med 1995;26(4):405-13.
10. Stiell IG, Wells GA, Hoag RH, Sivilotti ML, Cacciotti TF, Verbeek PR, et al. Implementation of the Ottawa Knee Rule for the use of radiography in acute knee injuries. Jama 1997;278(23):2075-9.
11. Oei EH, Nikken JJ, Verstijnen AC, Ginai AZ, Hunink MG. MR imaging of the menisci and cruciate ligaments: a systematic review. Radiology 2003;226(3):837-48.
12. Moher D, Schulz KF, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. Jama 2001;285(15):1987-91.
13. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. Ann Intern Med 2003;138(1):40-4.
14. Nikken JJ. MRI for the initial evaluation of acute wrist, knee, and ankle trauma. Rotterdam: Erasmus MC; 2003.
15. Hosmer DW, Lemeshow S. Assessing the fit of the model. In: Applied Logistic Regression. New York: J Wiley; 1989. p. 135-75.
16. Hunink MGM, Glasziou PP, Siegel JE, al e. Decision making in health and medicine: integrating evidence and values. Cambridge, England: Cambridge University Press; 2001.
17. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 1983;148(3):839-43.

18. Akaike H, editor. Information theory as an extension of the maximum likelihood principle. Budapest: Akademia Kiado; 1973.
19. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 1996;49(12):1373-9.
20. Vincken PW, ter Braak BP, van Erkell AR, de Rooy TP, Mallens WM, Post W, et al. Effectiveness of MR imaging in selection of patients for arthroscopy of the knee. Radiology 2002;223(3):739-46.
21. Kersting-Sommerhoff B, Gerhardt P, Golder W, Hof N, Riel KA, Helmberger H, et al. [MRI of the knee joint: first results of a comparison of 0,2-T specialized system and 1,5-T high field strength magnet]
MRT des Kniegelenks: Erste Ergebnisse eines Vergleichs von 0,2-T- Spezialsystem mit 1,5-T-Hochfeldmagnet. Rofo Fortschr Geb Rontgenstr Neuen Bildgeb Verfahr 1995;162(5):390- 5.
22. Vellet AD, Lee DH, Munk PL, Hewett L, Eliasziw M, Dunlavy S, et al. Anterior cruciate ligament tear: prospective evaluation of diagnostic accuracy of middle- and high-field-strength MR imaging at 1.5 and 0.5 T. Radiology 1995;197(3):826-30.
23. Kladny B, Gluckert K, Swoboda B, Beyer W, Weseloh G. Comparison of low-field (0.2 Tesla) and high-field (1.5 Tesla) magnetic resonance imaging of the knee joint. Arch Orthop Trauma Surg 1995;114(5):281-6.
24. Rutt BK, Lee DH. The impact of field strength on image quality in MRI. J Magn Reson Imaging 1996;6(1):57-62.
25. Barnett MJ. MR diagnosis of internal derangements of the knee: effect of field strength on efficacy. AJR Am J Roentgenol 1993;161(1):115-8.
26. Hayes CW, Brigido MK, Jamadar DA, Propeck T. Mechanism-based pattern approach to classification of complex injuries of the knee depicted at MR imaging. Radiographics 2000;20 Spec No:S121-34.
27. Sanders TG, Medynski MA, Feller JF, Lawhorn KW. Bone contusion patterns of the knee at MR imaging: footprint of the mechanism of injury. Radiographics 2000;20 Spec No:S135-51.
28. Boeree NR, Ackroyd CE. Assessment of the menisci and cruciate ligaments: an audit of clinical practice. Injury 1991;22(4):291-4.



Acute ankle trauma - the value of a short dedicated extremity MRI examination in predicting subsequent treatment.

Jeroen J. Nikken, MD [1,2], Edwin H.G. Oei, MSc [1,2], Abida Z. Ginai, MD, PhD [1], Gabriel P. Krestin, MD, PhD [1], Jan A.N. Verhaar, MD, PhD [3], Arie B. van Vugt, MD, PhD [4], Myriam G.M. Hunink, MD, PhD [1,2]

From the Program for the Assessment of Radiological Technology (ART Program). Department of Radiology [1], Department of Epidemiology and Biostatistics [2], Department of Orthopaedic Surgery [3] and Department of Traumatologic Surgery [4], Erasmus MC, University Medical Center, Rotterdam. the Netherlands.

Submitted for publication

Abstract

Purpose
To assess the value of a short MRI examination in addition to or instead of radiography in all patients presenting with acute ankle trauma, in order to identify patients who require additional treatment versus patients that will not need additional therapy and can be safely discharged from further follow-up.

Methods and materials
In a prospective controlled trial 197 patients with recent ankle trauma were randomized between radiography and the combination of radiography and a short MRI examination. Data were collected on sex, age, side, trauma mechanism, radiography result, MRI result, and treatment after the initial visit to the hospital. Using univariable and multivariable regression analysis, four models were created to predict treatment.

Results
Additional treatment was necessary in 109 of the 197 patients. In univariable analysis age (Odds ratio (OR) 1.02, 95% confidence interval (CI) 1.00-1.04), radiography result (OR 7.92, 95% CI 3.17-19.8) and positive or equivocal MRI result vs no MRI made (OR 2.42, 95% CI 1.25-4.70) predicted treatment. In the multivariable analysis a positive or equivocal MRI result (OR 2.61, 95% CI 1.28-5.30) contributed significantly to the prediction of subsequent treatment. However, a negative MRI result did not contribute significantly (OR 0.66, 95% CI 0.27-1.61).

Conclusion
A limited MRI examination in the initial evaluation of acute ankle injury in addition to radiography has additional predictive value in identifying patients who need treatment, but does not add significant information in identifying those who can be discharged from further follow-up.

Introduction

Traumatic injury of the ankle is a common reason for patients to visit an emergency department. The evaluation of ankle injury generally consists of taking a history and performing physical examination, frequently complemented with radiography. In most cases this strategy is reliable in the detection of fractures, but it is considerably less suitable to assess cartilage and ligamentous lesions (1, 2). Lateral stress radiography may give some information about lateral ankle ligamentous injury but in the acute phase pain often prevents a proper examination and the value of stress radiography is debated (3-5). A suitable imaging modality to demonstrate soft tissue lesions is MRI (5-8), but the high costs and the limited availability of most current (high field) MRI systems as well as the long duration of the examination have been a major hindrance to a wide application of MRI for acute ankle injury. This may, however, change with the advent of low field dedicated extremity MRI systems (9-11). Low field dedicated extremity MRI systems are considerably less expensive than high field systems and do not require costly and space occupying infrastructural changes. In addition, by their dedicated nature, they are only suitable for peripheral joint examination and thus not overbooked for other indications.

The potential value of MRI in the acute phase after trauma is in identifying patients who will not require any treatment and can thus be discharged without any further follow-up. Furthermore, patients that do have significant soft tissue injury are likely to be diagnosed at an earlier stage, avoiding delay in treatment, with potential cost saving to society because of diminished productivity loss. To our knowledge MRI has not been evaluated for this purpose before.

The purpose of this study was to assess the value of routine application of a short MRI examination after acute ankle injury in addition to or instead of radiography, in order to identify those patients who require treatment versus those that will not need additional therapy and can be safely discharged from further follow-up.

Materials and Methods

Study design

The data for this study were obtained from prospective randomized controlled clinical trial (RCT) analyzing the costs and effectiveness of a short MRI examination in all patients with acute ankle trauma and used to create prediction models. Although it is not necessary to perform an RCT to create prediction models, the data can be used for this purpose. Where applicable we used reporting standards for RCT's (12) and otherwise those for reporting diagnostic studies (13). Patients referred by a traumatologist or orthopaedic surgeon to the radiology department of

our university hospital for radiography of a recent ankle injury (within 7 days of trauma) were asked to participate in the study. Patients that did not undergo radiography were not included to ensure that patients with very mild trauma would not be enrolled. All patients were informed about the study by leaflet and by oral explanation. Informed consent was obtained from all participating patients. The study was approved by the institutional review board.
The included subjects were randomized to either plain radiography only or radiography combined with a short MRI examination. Randomization was carried out by drawing from consecutively numbered sealed envelopes containing computer generated random assignments. Patients were excluded if they had additional substantial injury of head, back, thorax or abdomen, if they had a compound ankle fracture, if they were in need of urgent treatment (e.g. in case of ankle luxation), if they had pre-existing complaints of the same ankle or if they were intoxicated by alcohol or drugs. Patients were included from 8.00 h to 23.00 h, 7 days a week. Inclusion was carried out by research staff and radiology technologists on service. Although inclusion was intended to be consecutive, only about half of all eligible patients was randomized.

TABLE 1. Parameters of the short MRI protocol

	Plane	TR (ms)	TE (ms)	Matrix	Slice thickness (mm)	Scan time (min:s)
(Half Fourier) SE T1-w	Sagittal	420-630	18-24	256x136	5.0	0:55
TME PD-w and T2-w	Axial	1980-2480	28-38/90	192x128	3.0	2:25
GE *	Coronal	400-420	15	192x152	5.0	0:44
STIR **	Coronal	1060	24-28	192x128	5.0	0:97

SE = spin echo, TME = turbo multi echo, PD = proton density, GE = gradient echo,
STIR = short-tau inversion recovery, TR = repetition time, TE = echo time.
For all sequences: field of view=200x140 mm, number of excitations=1
* Flip angle : 75 degrees
** Inversion time : 80-85 ms

Radiography was performed in anteroposterior and lateral direction. MRI was performed immediately following radiography on a 0.2 Tesla dedicated extremity MRI system (Artoscan M, Esaote , Genoa, Italy). A standard MRI examination of the ankle takes 30 to 45 minutes. This would be both too time consuming and too costly if all patients with acute ankle injury were to be examined.

Therefore we used a shortened MRI scanning protocol (table 1). This protocol was achieved by using only one excitation for each sequence, a rectangular field of view and by limiting the phase encoding direction of the matrix. Shortening of the duration of the examination comes at the cost of some loss in quality, but in a pilot study we found that the image quality was acceptable (14). The total acquisition time was on average 6:30 minutes with a total examination time, including MRI-system start-up and patient positioning, of approximately 15 minutes. A dual phased array ankle coil was used or, in case of a very large foot, a dual phased knee coil. An experienced musculoskeletal radiologist (A.Z.G) or a radiology resident immediately assessed the MRI without knowledge of the radiography result. The MRI result was communicated to the physician, so that the information was available for the decision on treatment or follow-up of the patient. If the MRI was assessed by the radiology resident (e.g. during weekends and in the evenings), the MRI was reassessed by the musculoskeletal radiologist the next day, without knowledge of the radiography result, and if the interpretation was different the treating physician was informed. The radiographs were initially assessed by the treating physician at the emergency ward, conform daily practice in our hospital. The next day the radiographs were reassessed by the radiologist in a plenary session with the trauma surgeon.

Outcome

As outcome measure we used additional treatment after initial presentation at our hospital. Since our goal was to differentiate patients that could be sent home without follow-up from patients in need of additional treatment, we did not include simple treatment measures at the first visit in this outcome (for example a pressure bandage). Additional treatment could consist of antipronation tape, softcast, plaster cast, physical therapy or surgery.

Data collection

Data were collected on age, sex, side of trauma, trauma mechanism (direct, indirect or unknown), radiography result, MRI result (traumatic injury visible, not visible or uncertain), and therapeutic procedures. We evaluated the effect of trauma mechanism on the outcome since indirect trauma is associated with rupture of ligaments and osteochondral lesions, whereas this is seldom the case in direct trauma. Trauma mechanism was scored as direct trauma if patients suffered a direct hit on the joint and as indirect trauma in case of a (combination of) inversion, eversion, hyperflexion or hyperextension injury. Data were obtained by direct assessment and by reviewing all information available from emergency department

records, hospital records, outpatient clinic records and the electronic hospital information system. Questionnaires were sent to all patients one week, six weeks, twelve weeks and six months after the initial visit to obtain information about treatment received outside our hospital. In case questionnaires were not returned, we interviewed the patient by telephone, provided they could be reached within two attempts of calling them. The final diagnosis was obtained by reviewing all information of diagnostic imaging, follow-up and questionnaires.
The follow-up period was 6 months. However, if the patient indicated on the questionnaire that he/she did not have daily complaints anymore the follow up was terminated, since no additional effects of the injury were to be expected.

Image Interpretation

The radiograph and the MRI were scored as 'positive' if signs of recent traumatic injury were visible. For the radiograph these signs were fracture, (sub)luxation, epiphysiolysis, osteochondral lesion, loose body with signs of recent traumatic origin (sharply edged), or widened distal tibio-fibular syndesmosis. The MRI was scored as positive if one or more of the following lesions were present: fracture, (sub)luxation, epiphysiolysis, osteochondral lesion, or ligamentous rupture. A ligament was scored as ruptured if discontinuity of the ligament was observed and the ligament was surrounded by edema or hematoma, distinguishing it from an old rupture. A partial ligament rupture was scored if the ligament was swollen but with intact continuity. If the result was uncertain this was scored as such. Since even minor traumatic injuries of the joint often cause abnormalities visualized by MRI, we excluded findings that generally do not need specific treatment. For this reason joint fluid, soft tissue edema, bone marrow edema, and hematoma without other lesions were not considered positive findings, although these findings may indicate the presence of another lesion.

Data analysis

Using univariable logistic regression analysis we determined the odds ratio and the 95% confidence interval of each variable in predicting the outcome. For the regression analysis we recoded the variable 'MRI result' in three possible outcomes: positive or uncertain, negative or no MRI performed. The MRI result was analyzed with the outcome 'no MRI performed' as reference. In this way we could use the whole study population in the analysis of the MRI result. When the number of uncertain MRI or radiography results was large enough to be analyzed as a separate dummy variable, this was done as such, otherwise the uncertain results were

grouped together with the positive results, since we wanted to avoid sending patients home without follow-up if the imaging result was uncertain.
Because the outcome of radiography and MRI result in the same patient is correlated, we created an additional integrated 'imaging variable' containing both radiography and MRI result.
This variable had five possible outcomes:

1) radiography negative and MRI not performed
2) radiography positive and MRI not performed
3) radiography and MRI both negative
4) radiography and MRI discrepant
5) radiography and MRI both positive.

With this combined 'imaging variable' we took into account the diagnostic interaction between radiography and MRI. The outcome 'radiography negative and MRI not performed' was considered the reference group for the combined 'imaging variable' analysis.
A variable was included in the multivariable logistic regression analysis if the chi square p-value was less than 0.1 in the univariable analysis. We combined the variables using multivariable logistic regression analysis to create a prediction rule for the identification of those patients that needed further treatment after the initial visit versus those requiring no additional treatment. Several models were evaluated to explore the predictive value of the imaging variables: in model 1 only radiography was added; in model 2 radiography was omitted and MRI was used instead, model 3 contained both radiography and MRI and in model 4 the combined 'imaging variable' was used.
The likelihood ratio test was used to assess the significance of each variable in the model in a stepwise backward approach. The calibration of the models was assessed using Hosmer and Lemeshow's Goodness of Fit Test (15). Calibration evaluates the degree of correspondence between the probabilities estimated by the model and the actual treatment received by the patients. If the p-value of the test statistic is large the model is well calibrated and the estimates fit the data well. The area under the ROC curve was used to evaluate the discriminative performance of the model and to compare predictive performance across models (16). The difference between the areas under the ROC curve of the different models was analyzed using the z-statistic and corrected for the fact that they were derived from the same cases. The correction factor was calculated using the Pearson product-moment correlation method (17).

Akaike's Information Criterion (AIC) was used to compare the predictive power of the models (18). The AIC is derived from the chi-square statistic and takes both the accuracy and the complexity of the model into account, by correcting for the number of degrees of freedom. In this way direct comparison of models is possible. An AIC equal to or smaller than 0 indicates no predictive value of the model. The significance of a difference in AIC of nested models (e.g. the model with radiography vs the model with radiography and MRI) was assessed using the likelihood ratio test. In case of non-nested models a difference in AIC larger than 2 was considered significant. All analyses were performed using SPSS for Windows (release 10.0.0; SPSS Inc., Chicago, IL).

Sample size

In multivariable logistic regression analysis at least 10 patients with an event and at least 10 patients without an event are necessary per variable analyzed (19). Because the data for this study were obtained from an RCT addressing the costs and effectiveness of early MR scanning of ankle trauma, the number of variables analyzed in the multivariable regression analysis was determined by the number of patients included in the RCT. Since 109 of the included patients received additional therapy and 98 patients did not receive additional therapy, we could analyze a maximum of 10 variables in the multivariable regression analysis.

Results

From August 1999 to May 2001, 202 patients were included, 107 males, 95 females with a mean age of 34.2 years, range 11.9-84.7 years. The number of patients randomized to radiography was 103 and to radiography followed by MRI 99. A flow diagram of subjects passing through the study is presented in figure 1.

One patient, randomized to MRI, did not undergo MRI examination because the MRI system was unavailable. Five randomized patients were incorrectly included in the study: three of them appeared to have pre-existing ankle complaints and two had a ruptured Achilles tendon; in fact, there was no indication for radiography in these last two cases. These 5 patients were excluded from the analysis. Of the remaining 197 patients 109 received treatment during the follow-up period. The final diagnoses of all patients are listed in table 2. Four patients with an ankle fracture were operated. Plaster immobilization was used in 40 cases of which 27 cases with a fracture, including one occult fracture of the medial malleolus, one occult fracture of the distal tibial plateau, one case with epiphysiolysis of the distal fibula, one case with combined fracture of the fibula and epiphysiolysis of the distal tibia, one case with an osteochondral lesion of the lateral talar dome, and 10 cases with lateral ankle ligamentous injury.

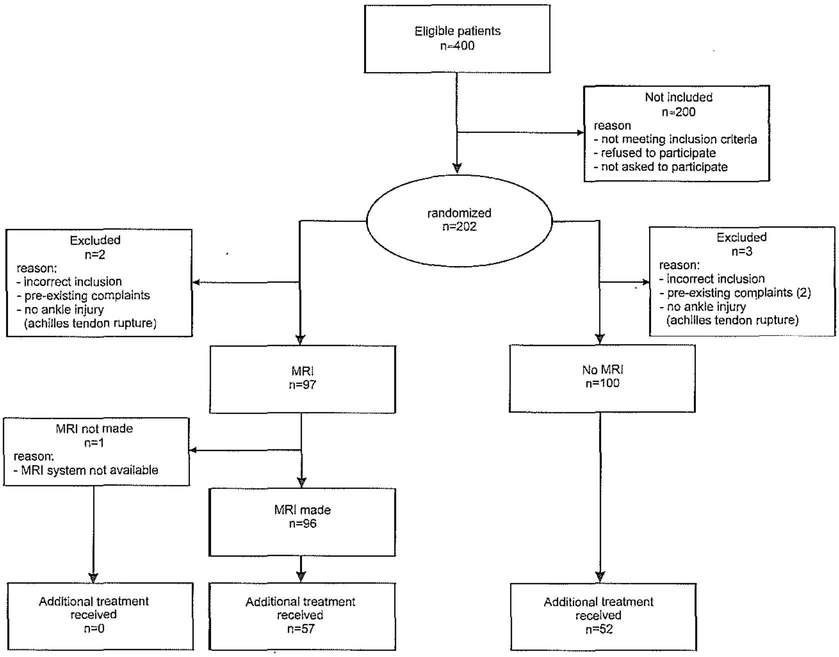


Figure 1. Flow diagram of subjects passing through the study.

Five fractures were missed on MRI, one distal fibula avulsion, one talar avulsion, one Weber B fracture and two metatarsal fractures that fell outside the field of view. Softcast or tape was used in 49 cases with lateral ankle ligamentous injury. In 14 cases one of the aforementioned treatments was combined with physical therapy, in 16 cases physical therapy was the only treatment. In 9 patients MRI revealed a rupture of the anterior tibio-fibular ligament; no combined rupture of the anterior and posterior tibio-fibular ligament was seen. Two patients in the control group underwent MRI examination later during follow-up. In one patient no abnormalities were found, in the other patient a talar dome fracture, already diagnosed on the initial radiograph, was evaluated. No patient was operated because of ankle instability during the follow-up period.

Regression analysis

In univariable analysis (see table 3) age, radiography result, and MRI result had a p-value smaller than 0.1 and were subsequently used in the multivariable logistic regression analysis. Sex, side, and trauma mechanism did not fulfill this criterion and were not used in the multivariable analysis.

Table 2. Final diagnosis in the experimental and the control group

Diagnosis	Experimental group (nr of lesions)	Control group (nr of lesions)
fracture	18	14
Weber A	1	3
Weber B	3	2
Weber C	1	0
medial malleolus	3	1
bimalleolar	2	2
distal tibia	1	0
distal tibia occult	1	0
navicular	1	1
talus	1	0
cuboid	0	2
cuboid + calcaneus	0	1
epiphysiolysis distal fibula	2	1
metatarsal	2	1
osteochondral lesion	3	1
haemarthrosis (haemophilia patient)	0	1
overstrained	0	1
ligament rupture*		
LTiFA	1	
LTiFP	0	
LTFA	28	
LTFA partial	6	
LTFP	0	
LTFA+LTiFA	8	
LTFA+LCF	1	
LTFA+LTiFA+LCF	2	
distortion †	32	78
contusion	3	4
unclear	0	1 §

LTiFA: anterior tibiofibular ligament, LtiFP: posterior tibiofibular ligament,
LTFA: anterior talofibular ligament, LTFP: posterior talofibular ligament,
LCF: fibulocalcaneal ligament

* In three cases the ligament injury was combined with a fracture and in two cases it was combined with an osteochondral lesion.

† A distortion was defined as an indirect ankle trauma, without signs of osseous injury and without proven ligament rupture.

§ One patient in the control group had persistent pain during the six months follow-up period, but no cause could be found.

Table 3. Univariable logistic regression analysis predicting the need for treatment after acute ankle injury - Odds ratios of covariates and p-value of the chi square test

Variable (predictor)	Proportion treated in group with predictor	Proportion treated in reference group	Univariable Odds ratio (95% CI)	p-value likelihood ratio test
Sex (male versus female)	56/105 (53%)	53/92 (58%)	0.84 (0.48-1.87)	0.55
Age (continuous)			1.02 (1.00-1.04)	0.08
Side (right versus left)	52/98 (53%)	57/99 (58%)	0.83 (0.48-1.46)	0.52
Indirect trauma (vs unknown or direct traumamechanism)	98/175 (56%)	11/22 (50%)	1.27 (0.52-3.09)	0.60
X-ray (pos or uncertain vs neg result)	40/46 (87%)	69/151 (46%)	7.92 (3.17-19.8)	<0.001
MRI result (vs MRI not made)				0.002
MRI positive or uncertain	47/66 (71%)	49/99 (49%)	2.42 (1.25-4.70)	
MRI negative	12/32 (38%)	49/99 (49%)	0.59 (0.26-1.33)	
Combined 'imaging variable' *				<0.001
radiography and MRI both negative	12/30 (40 %)	32/79 (41%)	0.96 (0.41-2.26)	
radiography positive, MRI not made	18/20 (90%)	32/79 (41%)	12.9 (2.80-59.7)	
radiography and MRI discrepant ‡	33/54 (61 %)	32/79 (41%)	2.26 (1.12-4.59)	
radiography and MRI both positive	14/14 (100%)	32/79 (41%)	∞ **	

Variables with a p-value < 0.10 were included in the multivariable analysis.

* Versus radiography negative and MRI not made as reference group.

‡ Discrepancy between radiography and MRI result: 3 times radiography positive and MRI negative, 51 times MRI positive and radiography negative.

** All individuals in this group received treatment, so the odds ratio is infinitely large.

The results of the multivariable logistic regression analysis are listed in table 4. In the multivariable analysis age was (borderline) significant. Both radiography and MRI result were significant predictors of treatment in all models, but for both imaging modalities only a positive result could predict the need for treatment, whereas a negative result could not predict the absence of this need (the confidence interval of the odds ratio of a negative result included unity). The strongest single predictor of treatment was a positive radiograph. The strongest category in the combined 'imaging variable' was the combination positive radiography and positive MRI result as all patients in this category received treatment. The combination 'negative radiography and negative MRI' could not predict that treatment was unnecessary.
The Hosmer-Lemeshow Goodness of Fit Test showed a non-significant p-value for all the models, although for models 1 and 3 the p-value was borderline. This indicates good calibration of models 2 and 4 and only moderate calibration of models 1 and 3. Akaike's Information Criterion of the model with radiography (model 1) was considerably larger than the AIC of the model with MRI (model 2), indicating a better predictive power of radiography. The combination of these variables in model 3 predicted the outcome significantly better than the model with only radiography, indicated by an increase in AIC (likelihood-ratio-test p-value 0.002). The difference in AIC between model 3 and model 4 was smaller than 2 and therefore considered as not significant. The correction for the interaction between radiography and MRI apparently did not increase predictive power of the model.
The area under the ROC curve was not significantly different between models 1, 3, and 4. However, the two models with both radiography and MRI and the combined 'imaging variable (models 3 and 4) showed a significantly larger area under the ROC curve than the model with only MRI (p-value 0.002 and 0.04 respectively), indicating better discriminatory power of the models with radiography and MRI compared to the model with only MRI.

Discussion

We performed a study to assess the value of a short dedicated extremity MRI examination in addition to or instead of radiography in predicting the need for treatment in patients presenting with acute ankle injury. Using logistic regression analysis four prediction models were evaluated. In the multivariable logistic regression analysis a positive MRI result in addition to radiography significantly added to predicting the need for treatment, but a negative MRI result had no significant predictive value.

Table 4. Multivariable prediction models predicting the need for treatment after acute ankle injury - odds ratios of the covariates and model performance statistics

	Model 1 Age and radiography (95 % CI)	Model 2 Age and MRI (95 % CI)	Model 3 Age, radiography and MRI (95 % CI)	Model 4 Age and the Combined 'imaging variable' (95 % CI)
Age	1.02 (1.00-1.04)	1.02 (1.00-1.04)	1.02 (1.00-1.04)	1.02 (1.00-1.04)
Radiography (versus no lesion)				
uncertain	4.34 (0.87-21.8)		9.92 (3.33-29.6)	
positive	4.23 (0.80-22.3)		9.71 (3.20-29.4)	
MRI result (versus no MRI made)				
negative	0.55 (0.24-1.26)		2.45 (1.26-4.79)	
positive or uncertain	0.66 (0.27-1.61)		2.61 (1.28-5.30)	
Combined 'imaging variable'*				
radiography and MRI both negative				0.95 (0.39-2.27)
radiography positive, MRI not made				13.40(2.88-62.4)
radiography and MRI discrepant				2.24 (1.10-4.58)
radiography and MRI both positive				∞**
Model performance statistics				
Hosmer-Lemeshow statistic	14.27 (p=0.08)	12.10 (p=0.15)	14.54 (p=0.07)	2.91 (p=0.94)
Akaike's Information Criterion	24.81	9.96	31.58	32.14
Area under ROC curve (95% CI)	0.70 (0.63-0.78)	0.64 (0.56-0.72)	0.74 (0.67-0.81)	0.72 (0.65-0.79)

* Versus radiography negative and MRI not made as reference group.
** All individuals in this group received treatment, so no odds ratio can be calculated.

Thus, although the MRI examination apparently demonstrates a number of significant lesions not shown on radiography and can identify patients who need additional treatment, it does not help in deciding which patient can be safely discharged from further follow-up.

In assessing the value of a short dedicated extremity MRI examination we did not aim at demonstrating the accuracy of the short MRI examination for all possible lesions. This would not be feasible because of the variety of possible lesions. Moreover, for many lesions it would not be possible to verify the diagnosis, since only surgery would be an acceptable reference standard. Only a minority of patients was treated surgically, so for most patients a suitable reference standard was not available. Instead, we chose to perform an outcome study approach and studied the predictive value of MRI for additional treatment after the initial visit, compared to the standard work-up.

Although MRI is very sensitive in the detection of occult fractures, its sensitivity in the detection of avulsion fractures is low. The limited field of view also caused MRI to miss two metatarsal fractures.

We were not absolutely sure if bone edema had to be regarded as 'no lesion requiring treatment' for the analysis. Bone edema as such does not need treatment, but it may add in predicting the need for treatment. Repeating the analysis with bone edema defined as an abnormal MRI result showed only a marginal change of the odds ratios in the models containing MRI (data not shown). One may argue that physical therapy should not be included as an endpoint, since the indication for physical therapy is often quite arbitrary. Repeating the analysis without physical therapy as an endpoint showed that age was not significant in the univariable analysis anymore, but the odds ratios of the other variables and the model performance statistics only changed marginally (data not shown).

Limitations of this study include the fact that, although inclusion was intended to be consecutive, only about half of all eligible patients was randomized. This was caused by patients not wanting to participate in the study, but also by eligible patients being missed. The main reason for missing patients was that inclusion took place from 8.00 - 23.00 hour, seven days a week and thus many different radiology technologists on service were supposed to recruit patients. Asking patients to participate in the study was forgotten regularly.

Another limitation was that we assumed that the decision to treat a patient was always correct. It is, however, possible that in some cases treatment was unnecessary, or that treatment was incorrectly withheld, although in the latter instance it is likely that most patients would return within six months because of persistent complaints and would have been treated at that time.

The follow-up period in our study was six months. Although it is likely that treatment will take place within 6 months after the injury, it is possible that some patients were operated after the follow-up period, for example for ankle instability.

However, most of these patients would have received physical therapy for proprioceptive training and improvement of muscle strength in the period preceding the operation and would have been coded as treated anyway.
The radiography result as a single predictor of therapy appeared to be of higher predictive value than the MRI result. This can be explained by the fact that the majority of ankle traumas are inversion injuries. In our hospital initial treatment of these injuries consists of bandage, followed by tape or softcast, depending on the amount of swelling and pain and does not depend on demonstration of ligament rupture. Ruptured ligaments are sometimes surgically repaired, but only in case of chronic ankle instability. Thus, in our hospital even if an MRI examination reveals one or more ruptured ligaments, the treatment is the same as for an ankle distortion without a ruptured ligament and the MRI will have limited additional value in predicting treatment. This is one of the reasons why we did not find a significant predictive value for negative imaging results. Another reason is that many patients will have physical therapy if pain persists irrespective of demonstrated lesions.
The short MRI protocol used in this study is associated with some loss of image quality, but in our view the subjective quality of the images we obtained was acceptable (Figure 2).

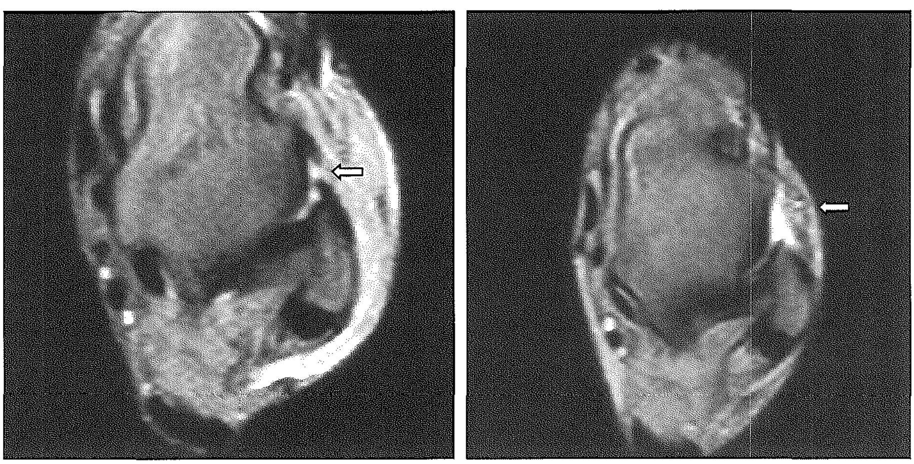


Figure 2a/b. a) Rupture of the anterior talo-fibular ligament with extensive haematoma in the overlying soft tissue. b) Extensive swelling of the anterior talo-fibular ligament indicating rupture of the ligament. Both images T2-weighted Turbo Multi Echo, FOV 200*140, matrix 256*128.

Compared to a state of the art protocol, however, it is possible that our accuracy was less, which may in part be responsible for the lack of discriminatory power that we found. The main reason, however, of lack of discriminatory power is the fact that besides ligamentous rupture very few other MR specific lesions were found in our patient group: we found no (osteo)chondral lesions, no osteochondrosis dissecans, only one occult fracture of the medial malleolus, and one epiphysiolysis of the distal fibula that was not recognized on the radiograph. By far the majority of traumatic ankle injuries consisted of fractures and ankle distortions, with or without ligament rupture.

If the treatment policy of an ankle distortion, with or without ligamentous rupture, consists of a pressure bandage followed by tape or softcast, the consequence of a false positive treatment is limited. This would be different if ruptured ankle ligaments are operated initially, in which case early information about the loss of integrity of the lateral ankle ligament is essential. Whether lateral ankle ligament ruptures should initially be treated conservatively or operatively is a matter of debate (20-23). A recent meta-analysis concludes that operative treatment leads to a better result than functional treatment (24). If the treatment policy of ruptured ankle ligaments were to shift in the direction of initial operation, there would be a role for an initial MRI examination in the acute ankle injury. Since this is not the case in our hospital, we were not able to evaluate the added value of early MR imaging in this respect.

Whether a short MRI examination should be routinely used in the evaluation of patients with acute ankle injury will mainly depend on the cost-effectiveness of this strategy, which we intend to evaluate in future work.

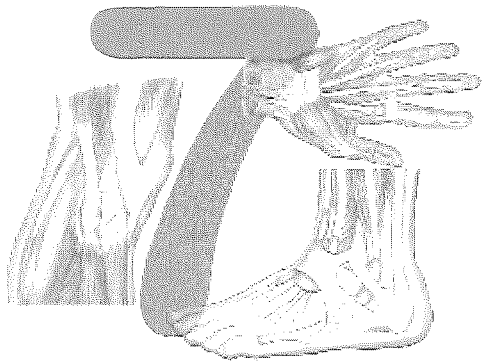
Conclusion

A limited MRI examination in the initial evaluation of acute ankle injury has additional predictive value over and above radiography in identifying patients who need treatment but cannot identify those who can be discharged from further follow-up. A limited MRI examination can not replace radiography to predict the need for additional treatment. MRI in the initial evaluation may, however, be valuable in a setting where ruptured ankle ligaments are immediately operated upon.

References

1. van Dijk CN. [CBO-guideline for diagnosis and treatment of the acute ankle injury. National organization for quality assurance in hospitals]
CBO-richtlijn voor diagnostiek en behandeling van het acute enkelletsel. Ned Tijdschr Geneeskd 1999;143(42):2097-101.
2. Stiell IG, Greenberg GH, McKnight RD, Nair RC, McDowell I, Reardon M, et al. Decision rules for the use of radiography in acute ankle injuries. Refinement and prospective validation. Jama 1993;269(9):1127-32.
3. Breitenseher MJ, Trattnig S, Kukla C, Gaebler C, Kaider A, Baldt MM, et al. MRI versus lateral stress radiography in acute lateral ankle ligament injuries. J Comput Assist Tomogr 1997;21(2):280-5.
4. Frost SC, Amendola A. Is stress radiography necessary in the diagnosis of acute or chronic ankle instability? Clin J Sport Med 1999;9(1):40-5.
5. Gaebler C, Kukla C, Breitenseher MJ, Nellas ZJ, Mittlboeck M, Trattnig S, et al. Diagnosis of lateral ankle ligament injuries. Comparison between talar tilt, MRI and operative findings in 112 athletes. Acta Orthop Scand 1997;68(3):286-90.
6. Kreitner KF, Ferber A, Grebe P, Runkel M, Berger S, Thelen M. Injuries of the lateral collateral ligaments of the ankle: assessment with MR imaging. Eur Radiol 1999;9(3):519-24.
7. Herber S, Kreitner KF, Kalden P, Low R, Berger S, Thelen M. [Low-field MRI of the ankle joint: initial experience in children and adolescents using an open 0.2 T MR-system] Niederfeld-MRT des Sprunggelenkes--Erste Erfahrungen bei Kindern und Jugendlichen an einem offenen 0,2 T MR-System. Rofo Fortschr Geb Rontgenstr Neuen Bildgeb Verfahr 2000;172(3):267- 73.
8. Breitenseher MJ. [Acute ankle injuries] Akutes Trauma des Sprunggelenks. Radiologe 1999;39(1):16-24.
9. Masciocchi C, Barile A, Satragno L. Musculoskeletal MRI: dedicated systems. Eur Radiol 2000;10(2):250-5.
10. Arbogast-Ravier S, Xu F, Choquet P, Brunot B, Constantinesco A. Dedicated low-field MRI: a promising low cost-technique. Med Biol Eng Comput 1995;33(5):735-9.
11. Hottya GA, Peterfy CG, Uffmann M, Hackl FO, LeHir P, Redei J, et al. Dedicated extremity MR imaging of the foot and ankle. Eur Radiol 2000;10(3):467-75.
12. Moher D, Schulz KF, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. Jama 2001;285(15):1987-91.
13. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. Ann Intern Med 2003;138(1):40-4.
14. Nikken JJ. MRI for the initial evaluation of acute wrist, knee, and ankle trauma. Rotterdam: Erasmus University; 2003.
15. Hosmer D, Lemeshow S. Applied logistic regression. New York: John Wiley & Sons, Inc.; 1989.

16. Hunink MGM, Glasziou PP, Siegel JE, al e. Decision making in health and medicine: integrating evidence and values. Cambridge, England: Cambridge University Press; 2001.
17. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 1983;148(3):839-43.
18. Akaike H, editor. Information theory as an extension of the maximum likelihood principle. Budapest: Akademia Kiado; 1973.
19. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 1996;49(12):1373-9.
20. Munk B, Holm-Christensen K, Lind T. Long-term outcome after ruptured lateral ankle ligaments. A prospective study of three different treatments in 79 patients with 11-year follow- up. Acta Orthop Scand 1995;66(5):452-4.
21. Agoropoulos Z, Papachristou G, Efstathopoulos N, Kokoroghiannis C. Late results of surgical repair in recent ruptures of the lateral ligament of the ankle. Injury 1997;28(8):531-4.
22. Kaikkonen A, Kannus P, Jarvinen M. Surgery versus functional treatment in ankle ligament tears. A prospective study. Clin Orthop 1996(326):194-202.
23. Kaikkonen A, Hyppanen E, Kannus P, Jarvinen M. Long-term functional outcome after primary repair of the lateral ligaments of the ankle. Am J Sports Med 1997;25(2):150-5.
24. Pijnenburg AC, Van Dijk CN, Bossuyt PM, Marti RK. Treatment of ruptures of the lateral ankle ligaments: a meta-analysis. J Bone Joint Surg Am 2000;82(6):761-73.



A Short Low-field MRI Examination in All Patients with Acute Peripheral Joint Injury
- Results of A Randomized Controlled Trial

Jeroen J. Nikken, MD [1,2], Edwin H.G. Oei, MSc [1,2], Abida Z. Ginai, MD, PhD [1], Gabriel P. Krestin, MD, PhD [1], Jan A.N. Verhaar, MD, PhD [3], Arie B. van Vugt, MD, PhD [4], M.G. Myriam Hunink, MD, PhD [1,2]

From the Program for the Assessment of Radiological Technology (ART Program).
Department of Radiology [1] and Department of Epidemiology and Biostatistics [2],
Department of Orthopaedic Surgery [3], Department of Traumatology [4],
Erasmus MC, University Medical Center Rotterdam.

Submitted for publication.

Abstract

Purpose:

To assess the costs and effectiveness of a short dedicated extremity MRI examination in acute wrist, knee, and ankle injury.

Methods and materials

Patients with recent trauma of wrist, knee, or ankle were randomized between radiography (reference group) and radiography followed by a short MRI examination (intervention group). Measures of effectiveness were quality of life, number of diagnostic procedures, time to last diagnostic procedure, number of days absent from work, and time to convalescence. A cost analysis was performed from the societal perspective.

Results

500 patients presenting with acute trauma of wrist, knee, or ankle were randomized. In the intervention group of knee-injured patients quality of life was significantly higher during the first six weeks (at one week p=0.003, at 6 weeks p=0.01), time to completion of the diagnostic workup was significantly shorter (mean 3.5 vs 17.3 days, p=0.003) and the number of diagnostic procedures was significantly lower (9 vs 35, p<0.001) compared to the intervention group. In patients with wrist or ankle injury the measures of effectiveness did not differ significantly between the intervention and the reference group. Knee-injured patients showed the largest difference in costs (intervention: 1820 Euro, reference: 2231 Euro), caused by a reduction in productivity loss. Costs were higher in the intervention group of wrist-injured patients and almost equal in ankle-injured patients. However, all cost differences were not significant.

Conclusion

A short MRI examination in acute wrist or ankle injury is neither cost saving nor effective. In knee-injured patients a short MRI examination shortens the time to completion of the diagnostic workup, reduces the number of diagnostic procedures, improves quality of life in the first 6 weeks, and may reduce costs associated with lost productivity.

Introduction

It is common practice to perform radiodiagnostic imaging of acute wrist, knee, or ankle injury in the initial stage. Most fractures can be detected fairly accurately using plain radiography. Plain radiography, however, provides limited information concerning soft tissue injury, such as a tendon or ligament rupture, a meniscus tear or a triangular fibrocartilage complex (TFCC) tear. Moreover, some fractures may be occult to plain radiography, a well-known example being the scaphoid fracture (1-3). Also, osteochondral lesions are sometimes not well visualized on plain radiography. To detect these lesions MRI would be the most suitable tool. MRI systems are, however, often not readily available, because they are in heavy demand by various medical disciplines. Furthermore, the high costs, the relatively long duration of the examination and the fact that about 5% of patients examined in whole body MRI systems cannot complete the examination because of claustrophobia (4) limit the use of MRI in the initial evaluation of acute joint injuries. Many of these constraints, however, may be overcome by the application of relatively inexpensive low field dedicated extremity MRI systems. Both purchase costs and maintenance costs are considerably lower compared to whole body systems and since only peripheral joints can be imaged, availability is less of a problem. Moreover, claustrophobia is almost ruled out, since only the arm or leg of interest is placed within the magnet bore, while the patient is lying on a chair next to the MRI system.

The aim of this study was to assess if the implementation of a short MRI examination performed on a low field dedicated extremity MRI system in addition to plain X-ray imaging is effective and cost saving compared to the current diagnostic imaging strategy (plain radiography only) in patients presenting with acute traumatic injury of wrist, knee, or ankle. Although the addition of an MRI examination to the initial diagnostic workup will obviously generate extra costs, the advantage of this extra examination may be expected in a shortening of the time to diagnosis. MRI examination at first presentation of the patient may save costs if it obviates the need of further follow-up, additional diagnostic procedures and temporary treatment measures, such as a plaster cast in case of a suspected scaphoid fracture. It may also save costs to society if the early diagnosis leads to an earlier treatment and hence to an earlier recovery, limiting loss of productivity. Since the costs of lost productivity can be considerable, limiting the number of days off work in a minority of patients may pay for the short MRI examination in all patients.

Methods

Study design

We conducted a prospective randomized controlled trial (RCT) at a university hospital. The study was approved by our institutional review board and was performed in accordance with the Helsinki Declaration of 1964, amended in 1989. We used the CONSORT statement to report the study results (5).Three parallel studies were conducted concerning the wrist, the knee, and the ankle. Patients were eligible if they presented with a recent injury of wrist, knee, or ankle (within 7 days of trauma) and were referred to the radiology department by a traumatologist or orthopaedic surgeon for radiography of the affected joint. After receiving written and oral information about the goal of the study, all participating patients gave informed consent. The subjects were randomized across two diagnostic strategies consisting of plain radiography alone (the reference group) and radiography followed by a short MRI examination (the intervention group). Randomization was achieved by drawing from consecutively numbered sealed envelopes containing computer generated random assignments. The randomization was stratified for joint and the envelopes were labeled accordingly. Patients were excluded if they had substantial injury of head, back, thorax, or abdomen, if they had a compound fracture, if they were in need of urgent treatment (e.g. in case of ankle luxation or open fracture), if they had pre-existing complaints of the same joint, or if they were intoxicated. Patients were included from 8.00 h to 23.00 h, 7 days a week, from August 1999 to May 2001. Inclusion was carried out by research staff and radiology technologists on service.

Sample size calculation

To demonstrate a mean difference in costs for each joint of Euro 500, with a standard deviation of the mean difference equal to 800, would require a sample size of 80 patients (power of 0.8, alpha of 0.05, two-sided test) per joint evaluated. Patients were enrolled until at least 80 subjects per joint were included .

MRI technique

Standard radiography was performed, if necessary supplemented by additional images. MRI was performed on a 0.2 Tesla dedicated extremity MRI system (Artoscan M, Esaote Biomedica, Genoa, Italy) immediately after radiography. We used a short MRI scanning protocol (Table 1) with one excitation for each sequence, using a rectangular field of view and a limited number of phase encoding steps.

Table 1. Parameters of the short MRI protocol

	TR (ms)	TE (ms)	Matrix	Slice thickness (mm)	Acquisition time (min:s)
(Half Fourier) SE T1-w	420-630	18-24	192-256x136	3.0-5.0	0:43-1.08
TME PD-w and T2-w	1980-2480	28-38/90	192-246x128	3.0	2:23-2:41
GE *	360-420	15	192x152	3.0-5.0	0:44-0:53
STIR ‡	830-1060	24-28	192x128	4.0-5.0	0:97-1:46

SE = spin echo, TME = turbo multi echo, PD = proton density, GE = gradient echo, STIR = short-tau inversion recovery, TR = repetition time, TE = echo time.
For all sequences: field of view=200x140 mm, number of excitations=1
* Flip angle = 75 degrees
‡ Inversion time = 80-85 ms

After a software upgrade in February 2000 the T1-w SE-half Fourier sequence was replaced by a T1-weighted Spin Echo sequence and for the TME sequence the first TE was shortened from 38 to 28 seconds (on the used MRI system the echo time is fixed and can not be changed by the user). The software upgrade included improved filtering and frequency sampling. The total acquisition time was 5-6 minutes for all joints, with a total examination time, including MRI system start-up and patient positioning, of approximately 15 minutes. An experienced musculoskeletal radiologist (A.Z.G.) or a radiology resident immediately assessed the MRI and reported the findings to the treating physician. If the MRI was assessed by the radiology resident (e.g. during weekends and in the evenings), the MRI was reassessed the next day by the musculoskeletal radiologist and if the interpretation was different the treating physician was informed. The radiographs were initially assessed by the treating physician at the emergency ward, conform daily practice in our hospital. The next day the radiographs were reassessed by the radiologist in a plenary session with the traumatologist.

Follow-up.
Clinical data were collected from patient hospital records and from the computerized hospital information system. The follow-up period was as long as the patient had daily complaints, as recorded on questionnaires, with a maximum of 6 months.

We expected this follow-up period of 6 months to be long enough to capture all relevant differences in the outcomes between the two strategies. Questionnaires were sent to the patients 1 week, 6 weeks, 3 months, and 6 months after inclusion. These questionnaires included quality of life instruments and questions about the number of days absent from work, number of days until the patient no longer had daily complaints, medical treatment in other institutions, treatment by the general practitioner, treatment by alternative medicine practitioners, and out-of-pocket expenses. If questionnaires were not returned, we interviewed the subjects by telephone and urged them to return the questionnaire, provided they could be reached within two attempts of calling them.

Measures of effectiveness

Measures of effectiveness were quality of life, time to completion of the diagnostic workup, number of additional diagnostic procedures, number of days absent from work and number of days to convalescence. Quality of life was measured with a descriptive instrument, the Short Form 36 Health Survey (SF-36) (6), and a valuative instrument providing general population values, the EuroQol (7, 8). The time to completion of the diagnostic workup was defined as the time from initial presentation to the date of last diagnostic imaging examination or diagnostic arthroscopy. The number of diagnostic procedures was assessed by reviewing the computerized hospital information system and by asking the subjects if any diagnostic procedure had been performed in another institution, using the questionnaires or by telephone inquiry. Likewise, the number of days absent from work and the number of days to convalescence (defined as the moment that the patient no longer had daily complaints) was asked in the questionnaires or, in case the questionnaire was not returned, by telephone inquiry.

Analysis of response

To check for response bias the baseline characteristics sex, age, and randomization result (MRI vs no MRI) of the subjects that returned questionnaires were compared with the baseline characteristics of the subjects that did not return any questionnaire. The same was done for the variables 'treatment' and 'number of days absent from work' as proxy measures of severity of disease, to assess if the response rate depended on severity of disease.

Analysis of effectiveness

The data were analyzed on an intention-to-diagnose-and-treat basis. The number of days to convalescence for both strategies was compared with a t-test.

The number of days to completion of the diagnostic workup, the number of diagnostic procedures, the number of days absent from work, and the evaluation of response bias were analyzed using an exact Wilcoxon-Mann-Whitney test and a chi-square test. Because of multiple testing we adjusted the p-value significance threshold of the tests from 0.05 to 0.01.

If a subject indicated on one of the four questionnaires that he or she was free of complaints, we expected no further change in quality of life caused by the initial trauma and EuroQol and SF-36 health survey data measured at that point in time were extrapolated to the remaining questionnaires. If at this stage less than 20 responses were available in the intervention group or in the reference group, stratified for joint, this was regarded as too few for meaningful analysis. If the EuroQol and SF-36 health survey data of a subject without complaints were not available for extrapolation (e.g. the patient told us by phone that he/she had no complaints anymore, but did not return the EuroQol and SF-36 questionnaires), mean values for the EuroQol and for the eight SF-36 domains were used to impute the censored health surveys. These mean values were derived from questionnaires of all complaint-free subjects that filled in the health survey questionnaires. Linear regression analysis was used to check if these mean values had to be adjusted for age and sex. Imputed values were only used in the analysis if they constituted less than 20% of the data. The preference based EuroQol tariff was calculated using the regression equation published by Dolan (9). Differences between the randomized groups in EuroQol and SF-36 domain scores were analyzed using a t-test.

Cost analysis

The cost analysis included all medical and non-medical costs relevant from the societal perspective associated with the initial injury during a follow-up period of 6 months. Direct medical costs included costs of hospital visits, diagnostic and therapeutic procedures, physical therapy inside or outside the hospital, visits to general practitioners, visits to alternative medical practitioners, and travelling costs for each visit. Costs of all diagnostic procedures were calculated using a bottom-up approach (10), taking into account the initial investment of equipment, additional costs during use, maintenance, years of use, discounting and annuitization (11), number of procedures per year, personnel costs, materials used, room rent, housekeeping, administration, and overhead costs. Costs were discounted at a rate of 3% per annum (12). Costs of initial treatment (e.g. bandage, plaster cast) were calculated likewise. Estimated actual costs of hospital visits and hospital admissions were obtained from the Dutch Council for Care Insurances (10).

For costs of operations reimbursement tariffs as established by the Dutch Central Organ for Tariffs in Healthcare were used. Indirect medical costs (medical costs in gained life years) were not considered, since we did not expect a difference in life years.
For the measurement of direct non-medical costs we took into account out-of-pocket expenses and patient time costs. Information on out-of-pocket expenses were derived from the questionnaires or from telephone inquiry. Patient time costs were determined by multiplying the time spent on medical procedures (such as hospital admissions, outpatient hospital visits, visits to the general practitioner, travel time, and waiting time) by the average net income of subjects (both working and not working) comparable in age for the year 2000 (source: Dutch Central Bureau for Statistics) (12) .
For indirect non-medical costs we assessed costs associated with production losses. Production losses were estimated using the 'friction cost method' as described by Koopmanschap et al and recommended in guidelines for cost-effectiveness analysis (12, 13). This method assumes that production losses are only generated during the time it takes to replace a sick employee - the friction period. Short term absence from work may cause limited loss of production, because work may be postponed and performed on return, or work may be taken over by colleagues. Long term absence from work generates costs by actual production losses, extra costs to maintain production, and costs of filling a vacancy and training a new worker in case of permanent replacement of a worker. Friction costs were calculated using friction cost data estimated for working people in the Netherlands in 1998, based on age and sex by Oostenbrink et al (10). These data were adjusted for general increase in income in 2000 compared to 1998. The friction period for the year 1998 was estimated at 4 months (10), but because of increased shortage on the labor market in 2000 we estimated the friction period to be 6 months (personal communication Koopmanschap). Costs were analyzed on an intention-to-diagnose-and-treat basis and compared with a t-test. We performed the cost analysis for the three joints separately.
Sensitivity analysis was used to study the robustness and generalizability of the study results. One-way sensitivity analysis was conducted by exploring a range from 50% to 200% of the baseline value for each single cost price, the time assigned to each medical procedure as well as travel time, friction costs, and friction period. Three-way sensitivity analysis was conducted on the parameters that were most sensitive in the one-way analysis.

All analyses were performed using Microsoft Excel 97 SR-2 (Microsoft Corp, Redmond, WA), SPSS for Windows (release 10.0.0; SPSS Inc., Chicago, IL, USA) and StatXact-4 for Windows (Cytel Software Cooperation, Cambridge, MA, USA) software packages.

Results

From August 1999 to May 2001 a total of 500 patients was included, 207 female and 293 male, mean age 34.8 years, range 12.0-85.0 years, standard deviation 15.1. A flow diagram of patients passing through the study is presented in figure 1.

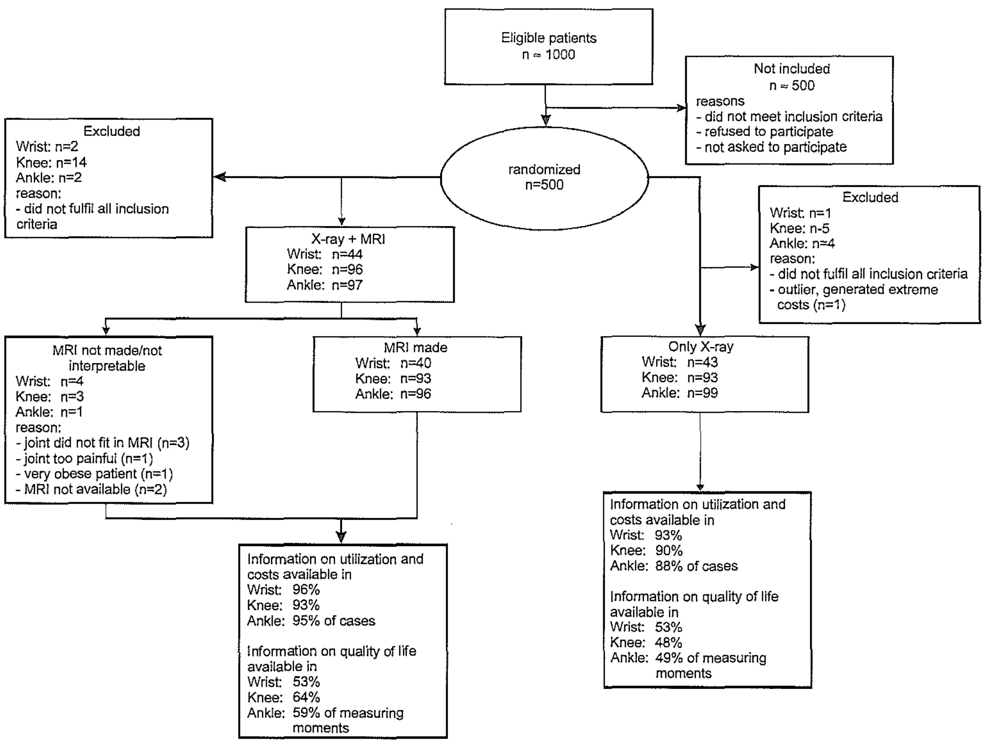


Figure 1. Flow diagram of subjects passing through the study.

Although inclusion was intended to be consecutive, only about half of all eligible patients was randomized. This could be attributed partly to patients refusing to participate in the study, but also to eligible patients not being asked to participate. This was caused by the fact that patients were included seven days a week, from 8.00 - 23.00 hour and thus many different radiology technologists on service were

supposed to ask patients to participate. Asking patients to participate in the study was forgotten regularly.
Twenty-eight randomized patients were excluded at some point in time after inclusion: on review 26 patients did not fulfill the inclusion criteria, one patient was excluded because there was strong suspicion of psychopathology with automutilation and one excluded patient was an outlier in several respects. This last subject was a 79 year old lady with a severe luxation fracture of the ankle who came to the hospital only 6 days after the injury. For logistic reasons she was transported to another hospital, where she was operated and hospitalized for more than 6 months, because she had many complications. Her medical costs were almost twice the total medical costs of all other patients with ankle injury together, and her time costs were almost 3 times the total time costs for all patients with ankle injury together.
Of the remaining 472 patients, 237 patients were allocated to the MRI strategy and 235 patients to the reference strategy. In 8 patients, allocated to the MRI strategy, the MRI was not made or not interpretable. In one case there was a technical failure, in one case the examination was interrupted because of pain, in one very obese patient the image quality was too poor for interpretation and in two cases the MRI system was not available. In three cases the joint of interest could not be positioned in the center of the magnet bore: one patient was not able to extend the elbow, one patient had a locked knee, and in one case the knee was too big to fit into the magnet. In the analysis these eight subjects stayed in the strategy as originally assigned.

Response characteristics

We found no significant difference in sex, age or number of days to resumption of work between patients that returned questionnaires and patients that did not return questionnaires. However, in ankle-injured patients the number of subjects that returned one or more questionnaires was slightly higher in patients who underwent MRI examination (84% vs 72%; $p=0.05$). Furthermore, in patients with knee and ankle injury this response rate differed significantly between patients who had undergone additional treatment vs those who did not return for additional treatment (knee: 93% vs 61%; $p<0.01$, ankle: 85% vs 68%; $p=0.01$).

Effectiveness

After extrapolation of EuroQol and SF-36 data in censored patients, data were only used for analysis if the number of complete questionnaires exceeded 20 in both the intervention group and the reference group stratified by joint. This was the case in

all groups except for the third and fourth questionnaire in patients with wrist injury. Linear regression analysis showed that mean EuroQol and SF-36-domain scores of patients without complaints were not significantly influenced by age or sex, and therefore total mean values of EuroQol and SF-36-domain values were used for imputation. In all four questionnaires, stratified for joint and randomization result, less than 20% of the questionnaires was imputed.
In patients with wrist and ankle injury the EuroQol tariff measured at one week, 6 weeks, 3 months, and 6 months after the injury was at no point in time significantly different between the intervention group and the reference group (Figure 2).

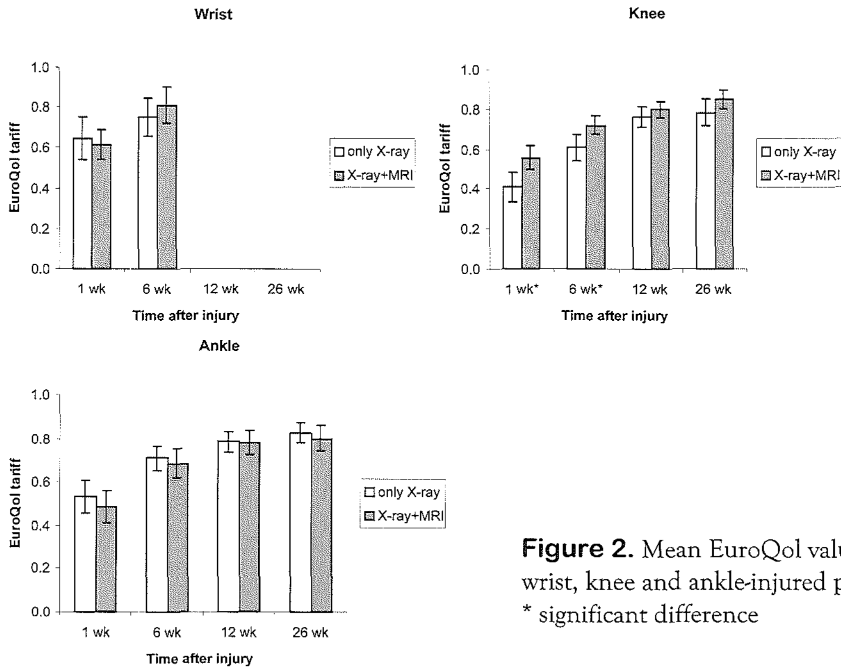


Figure 2. Mean EuroQol values for wrist, knee and ankle-injured patients.
* significant difference

In knee-injured patients the EuroQol tariff at one week and six weeks was significantly higher in the intervention group (t-test p-value 0.003 resp 0.01). Three months after the injury this difference was not significant anymore.
All the SF-36 domains (except general health) demonstrated improvement in quality of life over time (Figure 3).

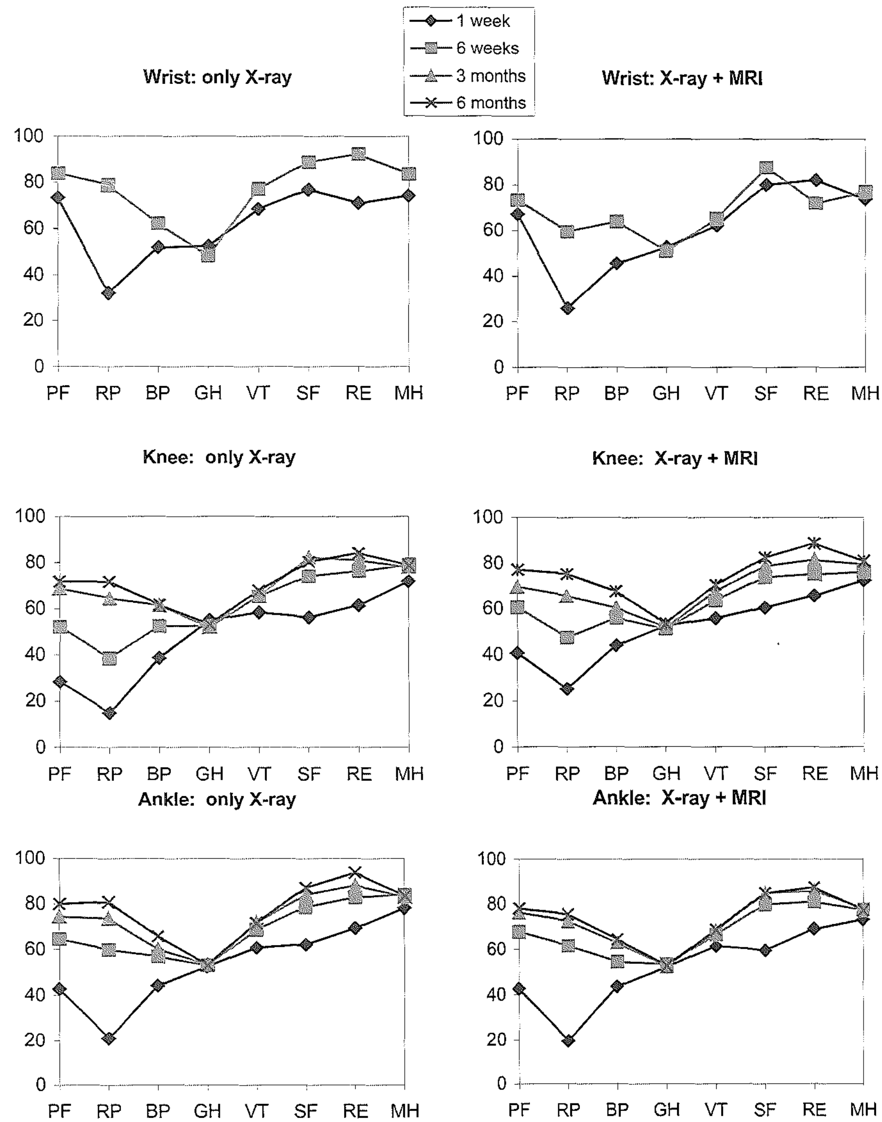


Figure 3. Mean SF-36 domain scores for wrist, knee, and ankle-injured patients at 1week, 6 weeks, 3 months and 6 months. For the wrist 3 months and 6 months scores were omitted because the paucity of data precluded meaningful analysis. PF= physical functioning, RP=role-physical, BP=bodily pain, GH=general health, VT=vitality, SF=social functioning, RE=role-emotional, MH=mental health. All domains (except general health) improved over time. The lines between the points are only to improve visual comparability and do not have intrinsic meaning.

There were, however, only a few significant differences between the intervention group and the reference group. For the wrist the vitality and role-emotional domains scored higher in the intervention group at 6 weeks. For the knee the physical functioning score was higher in the intervention group at one week. For the ankle the mental health score was lower in the intervention group at 6 weeks and at 6 months.
Only in the intervention group of knee-injured patients the number of diagnostic procedures performed during follow-up and the time to completion of the diagnostic workup were significantly lower in the intervention group of patients with knee injury (p-value 0.01) (Table 2).
The duration of absence from work did not differ significantly between any of the intervention and reference groups. Although for all joints, time to convalescence was lower in the intervention group, the difference compared to the reference group was not significant.

Table 2. Comparison of measures of effectiveness across strategies

	Wrist			Knee			Ankle		
	Ref. (n=43)	Int. (n=44)	p-value	Ref. (n=93)	Int. (n=96)	p-value	Ref. (n=99)	Int. (n=97)	p-value
Number of additional diagnostic procedures*	6	3	0.41	35	9	<0.001	10	1	0.04
Mean time to last diagnostic procedure (days)*	0.98	4.3	0.43	17.3	3.5	0.003	3.4	0.07	0.10
Mean duration of absence from work (days)*	5.0	8.6	0.31	12.4	9.6	0.36	12.6	11.2	0.70
Mean time to convalescence (days) ‡	54.3	49.6	0.71	76.6	66.2	0.45	55.9	54.0	0.84

Ref.: only radiography, the reference group
Int.: radiography + MRI, the intervention group
* difference tested with Wilcoxon-Mann-Whitney exact test
‡ difference tested with t-test

Costs

The mean total medical and non-medical costs associated with the joint injury are listed in table 3, and graphically displayed in figure 4.

Table 3. Mean costs associated with the initial injury from the societal perspective (costs in Euro)

	Wrist				Knee				Ankle			
	Ref.	Int.	Diff.	T-test p-value	Ref.	Int.	Diff.	T-test p-value	Ref.	Int.	Diff.	T-test p-value
Medical costs	344	306	-38	0.59	606	691	85	0.51	328	368	40	0.48
Travel costs	12	10	-2	0.32	16	17	1	0.52	12	11	-1	0.29
Friction costs	519	766	247	0.42	1510	993	-517	0.18	1285	1265	-20	0.96
Time costs	31	24	-7	0.53	99	118	19	0.70	35	30	-5	0.73
Total costs	905	1106	201	0.53	2231	1820	-411	0.35	1660	1675	-15	0.98

Ref.: only radiography, the reference group
Int.: radiography + MRI, the intervention group
Diff.: difference in costs between the intervention and the reference groups; a negative sign indicates a reduction of costs with the use of MRI.

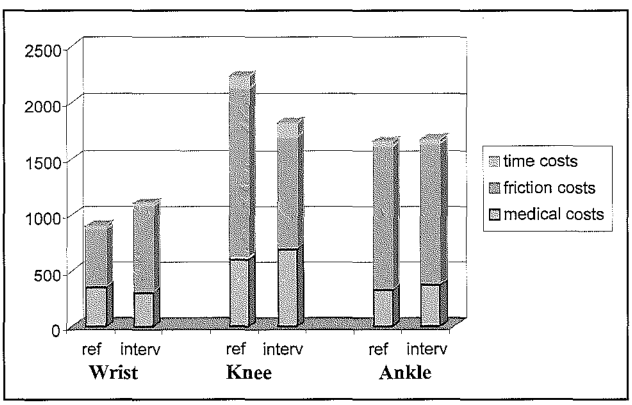


Figure 4. Visual representation of the costs described in table 3.
Travel costs were omitted in the figure since these costs were too low to depict.
ref = only radiography, the reference group
interv = radiography + MRI, the intervention group.

The largest difference in total costs was found in the patients with knee trauma. This difference was attributable to a reduction in the friction costs with the use of MRI. In patients with wrist injury the friction costs increased slightly with the use of MRI. In ankle-injured patients all costs were quite similar in the intervention and the reference group. For all three joints the differences in costs were not significant.

Sensitivity analysis

One-way sensitivity analysis in wrist and knee-injured patients showed that the friction costs was by far the most sensitive single variable on total cost difference between the intervention and reference group (Table 4).

Table 4. One-way sensitivity analysis of the mean total cost difference between the intervention and the reference strategy.

The costs difference between the two strategies was recalculated for each single variable at 50% and 200% of the original value of that variable to assess the influence of variation of that variable on the total cost difference. A negative cost difference indicates a reduction in costs with the use of MRI. The baseline cost difference is the cost difference between the strategies with all variables fixed at 100%. The range is a measure of sensitivity of the cost difference to that variable. Only the 11 most sensitive variables are presented. Costs are in Euro.

	Wrist			Knee			Ankle		
	50 %	200 %	range	50 %	200 %	range	50 %	200 %	range
Baseline cost difference between the two strategies	201			-412			14		
Friction costs	77	448	371	-153	-929	776	24	-5	29
Friction period	165	201	36	-264	-412	148	133	14	119
Short MRI examination	172	257	85	-440	-355	85	-14	71	85
Outpatient visit	222	159	63	-425	-385	40	21	2	19
Time costs	204	194	10	-421	-392	29	17	10	7
Hospital day	208	185	23	-418	-398	20	14	15	1
Physiotherapy	200	201	1	-416	-402	14	22	-1	23
Standard MRI examination	201	201	0	-408	-418	10	16	14	2
Radiograph	214	173	41	-413	-408	5	11	22	11
Operative therapy	208	185	23	-410	-415	5	9	25	16
General practitioner visit	195	211	16	-411	-412	1	12	19	7

In ankle-injured patients friction costs were very similar in the intervention group and the reference group, so variation did not influence the difference between the groups as much as in the other two joints. In knee and ankle injury the duration of the friction period was also influential. Costs of the short MRI examination was the third most sensitive variable in all groups. Three-way sensitivity analysis of friction costs, friction period, and short MRI examination costs (not shown) demonstrated that within plausible ranges of the values for the analyzed variables, the conclusions for the wrist and the knee did not change. For ankle-injured patients the results were sensitive to the friction period.

Discussion

Only few cost analysis studies have been published on the application of MRI in joint trauma. Most of these studies were performed in the UK and consider the National Health Service perspective (14-16). To our knowledge no studies have been published assessing the cost-effectiveness of MRI in acute joint trauma from the societal perspective.

We expected the main benefit of early diagnosis in acute trauma of wrist, knee, or ankle using a dedicated-extremity-MRI system to be twofold: a decreased need for follow-up and a reduction in time to convalescence. The decreased need for follow-up could lead to potential cost-savings by a reduction in follow-up visits, additional imaging studies, diagnostic arthroscopy, and temporary treatment measures. This would also reduce patient time costs, travel costs, and friction costs. A reduction in the time to convalescence would result from a reduction in the time to diagnosis, with the associated earlier treatment. This could lead to cost-savings to society due to an earlier resumption of work (i.e. a reduction in friction costs). The aforementioned effects could lead to a total cost reduction, provided that the cost-savings outweigh the additional costs of performing a short dedicated-extremity-MRI examination in every patient.

In our patient population we could not demonstrate that a short MRI examination in addition to radiography was cost saving from the societal perspective. In knee-injured patients mean total costs were, however, substantially lower in the intervention group due to reduced costs of lost productivity, but the difference was not significant. The influence of a short MRI examination on other outcomes was demonstrated only in knee-injured patients: quality of life was improved, the number of additional examinations was lower in the intervention group, and the time to diagnosis was reduced by 12 days. Both duration of absence from work as

well as time to convalescence did not differ significantly between the intervention group and the reference group for all three joints.
A limitation of this study is the setting in which it was performed. The Dutch health care system is facing long and variable waiting lists. The variation in waiting lists for outpatient visits and surgical procedures may have influenced the time to convalescence as outcome measure. In our hospital the waiting list for arthroscopy at the time of the trial was 5 months. We therefore used the time to the last diagnostic procedure as a more representative measure. Likewise it is plausible that the friction costs were also influenced by the waiting lists. Although the extra friction costs caused by waiting lists are incurred in both the intervention group and the reference group, the increase of these costs combined with variation could have masked a difference between the groups.
We recognize that early detection and (surgical) treatment of a condition is not always necessary for recovery. Some meniscal tears, TFCC lesions or ligament injuries may heal without specific treatment (17). In this perspective the presence of waiting lists is not solely unfavorable for the patient, but may also allow time for the natural recovery of a condition.

The response rate on the questionnaires was low for all groups. At the moment of inclusion the importance of filling in the questionnaires was stressed to the patients and they were urged to return them. If the questionnaires were not returned in time the patients were called by phone to remind them. In spite of our efforts, the response rate was less than optimal. Typical response rates for postal questionnaires are about 60-65% (18-27). A reason for the lower response rates in our study may be that patients are less willing to take the effort to respond if the disease does not have a major effect on their health and the disease is not likely to last long. Furthermore, about 20% of our patients were non-natives, with a varying ability to speak the national language. We did not record how well a subject was able to speak the national language, but subjects with a non-native name showed a significantly lower response rate than subjects with a native name (chi square p-value 0.001), which may be attributable to their knowledge of the language, or possibly by cultural differences. The subjects with non-native names were evenly distributed among the two randomized groups (chi square p-value 0.23), so their relatively large contribution to the high non-response rate is not likely to have caused response bias in the difference between the groups. For the analysis of costs we were much less dependent on the low response rate, since we could retrieve most data from the hospital computer system and patient records, complemented by telephone inquiry.

The addition of MRI to the initial diagnostic evaluation had no influence on the quality of life from three to six months after the injury as measured with the EuroQol and SF-36 instruments. The short term (1-6 weeks) EuroQol outcome in knee-injured patients was, however, significantly higher in the intervention group compared to the reference group. This may reflect a more adequate initial treatment in the intervention group. Another possible explanation would be that despite randomization the intervention group had less severe injury than the reference group. There were, however, no indications of a difference in severity of injury between the two groups: there was no significant difference in number of patients receiving treatment, number of hospital admissions or number of hospital days between the two groups.
In this pragmatic RCT no blinding was used. Patients knew whether an MRI examination was carried out or not. This influenced the response rate of patients with ankle injury: patients who had undergone MRI were more likely to return their questionnaires. It potentially also influenced the time to convalescence, since patients may have been reassured by a negative MRI result. Reassurance is, however, part and parcel of performing the MRI and should not be regarded as a bias in this context. Blinding could have been achieved by performing a placebo-MRI, or by not giving the information of the MRI (28). Because of the expected confusion and opposition of patients with an acute joint injury to such a study design, as well as ethical considerations of not giving the information, we chose a non-blinded pragmatic study design.
Part of the pragmatic design of the study was the MRI system software upgrade during the study, which may have improved the quality of the images in the latter part of the study. Since MRI software upgrades can be expected in daily practice we considered the upgrade acceptable in this pragmatic study. The extent of acceptance of the short MRI examination by clinicians may also have introduced a bias. We noticed that clinicians were sometimes hesitant to use the results of the short MRI examination in their decisions. This 'acceptance curve' of the new technology may have biased the outcome of the study in favor of the reference strategy. Nevertheless, we did find that if a short MRI examination had been performed in the acute setting, patients were less likely to undergo further examinations during follow-up.

In this study a short MRI scanning protocol was used to make the examination time acceptable for routine use at the emergency department as well as to reduce costs. Shortening of the duration of the MRI examination comes at the cost of some loss in quality, but in a pilot study we found that the image quality was acceptable to detect most significant lesions. It is likely that using this short MRI protocol some diagnoses will be missed, that would have been detected if a standard scanning protocol would have been used. It is therefore important to realize that the goal of the studied short MRI examination was not to replace standard MRI examination, but to diminish diagnostic uncertainty at the initial presentation of the patient, resulting in a decrease in total costs with a potential gain in effectiveness. In case the lower quality of the short protocol images would result in a wrong diagnosis, this would likely lead to an increase in costs and a decrease in effects, including a decrease in quality of life. However, if the overall effect of an initial diagnostic strategy including routine short MRI examination results in an overall decrease in costs, without loss of effectiveness, or even an increase in effectiveness, the new strategy is worthwhile considering.

Although we could not demonstrate with statistical significance that the application of a short MRI in all patients with knee injury is cost-saving from the societal perspective, our results suggest that this might well be the case. For the wrist the strategy including the short MRI examination increased the total costs, without a change in effectiveness. For the ankle the application of MRI did not change the total costs and had hardly any influence on the effects.

Conclusion

A short MRI examination in all patients with wrist or ankle injury in addition to radiography is neither cost saving nor effective. Our results do, however, suggest that a short MRI examination in the initial workup of patients with acute knee injury shortens the time to completion of the diagnostic workup, reduces the number of additional diagnostic procedures required, and may reduce overall costs by reducing the costs associated with lost productivity.

References

1. Bretlau T, Christensen OM, Edstrom P, Thomsen HS, Lausten GS. Diagnosis of scaphoid fracture and dedicated extremity MRI. Acta Orthop Scand 1999;70(5):504-8.
2. Gabler C, Kukla C, Breitenseher MJ, Trattnig S, Vecsei V. Diagnosis of occult scaphoid fractures and other wrist injuries. Are repeated clinical examinations and plain radiographs still state of the art? Langenbecks Arch Surg 2001;386(2):150-4.
3. Breitenseher MJ, Metz VM, Gilula LA, Gaebler C, Kukla C, Fleischmann D, et al. Radiographically occult scaphoid fractures: value of MR imaging in detection. Radiology 1997;203(1):245-50.
4. Moss ML, Buongiorno PA, Clancy VA. Intranasal midazolam for claustrophobia in MRI. J Comput Assist Tomogr 1993;17(6):991-2.

5. Moher D, Schulz KF, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. Jama 2001;285(15):1987-91.
6. Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992;30(6):473-83.
7. EuroQol--a new facility for the measurement of health-related quality of life. The EuroQol Group. Health Policy 1990;16(3):199-208.
8. Brooks R. EuroQol: the current state of play. Health Policy 1996;37(1):53-72.
9. Dolan P. Modeling valuations for EuroQol health states. Med Care 1997;35(11):1095-108.
10. Oostenbrink JB, Koopmanschap MA, Rutten FFH. Handleiding voor kostenonderzoek, methoden en richtlijnprijzen voor economische evaluaties in de gezondheidszorg. Amstelveen: College voor zorgverzekeringen; 2000.
11. Hunink MGM, Glasziou PP. Decision making in health and medicine - integrating evidence and values. Cambrige: Cambridge University Press; 2001.
12. Gold MR, Siegel JE, Russell LB, Weinstein MC. Cost-effectiveness in health and medicine. Oxford: Oxford University Press; 1996.
13. Koopmanschap MA, Rutten FF, van Ineveld BM, van Roijen L. The friction cost method for measuring indirect costs of disease. J Health Econ 1995;14(2):171-89.
14. Bryan S, Weatherburn G, Bungay H, Hatrick C, Salas C, Parry D, et al. The cost-effectiveness of magnetic resonance imaging for investigation of the knee joint. Health Technol Assess 2001;5(27):1-95.
15. Dorsay TA, Major NM, Helms CA. Cost-effectiveness of immediate MR imaging versus traditional follow-up for revealing radiographically occult scaphoid fractures. AJR Am J Roentgenol 2001;177(6):1257-63.
16. Raby N. Magnetic resonance imaging of suspected scaphoid fractures using a low field dedicated extremity MR system. Clin Radiol 2001;56(4):316-20.
17. Solomon DH, Simel DL, Bates DW, Katz JN, Schaffer JL. The rational clinical examination. Does this patient have a torn meniscus or ligament of the knee? Value of the physical examination. Jama 2001;286(13):1610-20.
18. Becker A, Seitz R, Jacobi E, Leidl R. [Cost evaluation by a patient questionnaire: pilot study of a weekly cost diary] Kostenmessung durch Patientenbefragung: Pilotstudie zu einem Kostenwochenbuch. Rehabilitation (Stuttg) 2001;40(1):12-20.
19. Biering-Sorensen F, Biering-Sorensen M. Sleep disturbances in the spinal cord injured: an epidemiological questionnaire investigation, including a normal population. Spinal Cord 2001;39(10):505-13.
20. Heath AC, Howells W, Kirk KM, Madden PA, Bucholz KK, Nelson EC, et al. Predictors of non-response to a questionnaire survey of a volunteer twin panel: findings from the Australian 1989 twin cohort. Twin Res 2001;4(2):73-80.
21. Bakker E, Van Gool JD, Van Sprundel M, Van Der Auwera C, Wyndaele JJ. Results of a questionnaire evaluating the effects of different methods of toilet training on achieving bladder control. BJU Int 2002;90(4):456-61.

22. Hanneman MJ, Sprangers MA, De Mik EL, Ernest van Heurn LW, De Langen ZJ, Looyaard N, et al. Quality of life in patients with anorectal malformation or Hirschsprung's disease: development of a disease-specific questionnaire. Dis Colon Rectum 2001;44(11):1650-60.
23. Campbell MJ, Waters WE. Does anonymity increase response rate in postal questionnaire surveys about sensitive subjects? A randomised trial. J Epidemiol Community Health 1990;44(1):75-6.
24. Balen AH, Fleming C, Robinson A. Health needs of adolescents in secondary gynaecological care: results of a questionnaire survey and a review of current issues. Hum Fertil (Camb) 2002;5(3):127-32.
25. Lund E, Gram IT. Response rate according to title and length of questionnaire. Scand J Soc Med 1998;26(2):154-60.
26. Duffy DL, Martin NG. Increasing the response rate to a mailed questionnaire by including more stamps on the return envelope: a cotwin control study. Twin Res 2001;4(2):71-2.
27. Maconochie N, Doyle P, Davies G, Lewis S, Prior S, Sampson P, et al. The study of reproductive outcome and the health of offspring of UK veterans of the Gulf war: methods and description of the study population. BMC Public Health 2003;3(1):4.
28. Hunink MG, Krestin GP. Study design for concurrent development, assessment, and implementation of new diagnostic imaging technology. Radiology 2002;222(3):604-14.

8

Costs and Effects of a Short Low-field MRI Examination in Acute Knee Injury

Jeroen J. Nikken, MD [1,2], Edwin H.G. Oei, MSc [1,2], Abida Z. Ginai, MD, PhD [1], Gabriel P. Krestin, MD, PhD [1], Jan A.N. Verhaar, MD, PhD [3], Arie B. van Vugt, MD, PhD [4], Myriam G.M. Hunink, MD, PhD [1,2]

From the Program for the Assessment of Radiological Technology (ART Program).
Department of Radiology [1], Department of Epidemiology and Biostatistics [2], Department of Orthopaedic Surgery [3] and Department of Traumatology [4], Erasmus MC, University Medical Center, Rotterdam. the Netherlands.

Submitted for publication.

Abstract

Purpose

To assess the costs and effectiveness of performing a short MRI examination following radiography in the evaluation of recent knee trauma.

Methods and materials

Subjects with recent knee injury were randomized between radiography and radiography followed by a short MRI examination. A third modeled strategy consisted of radiography in all patients followed by MRI if the radiograph showed no fracture. Data were collected on quality of life, time to completion of the diagnostic workup, number of additional examinations, time absent from work, time to convalescence and medical and non-medical costs. A cost analysis was performed from the societal perspective.

Results

208 patients were randomized. At 1 and 6 weeks quality of life was lowest if only radiography had been performed. Quality of life was similar across the two MRI strategies. Likewise, the time to completion of the diagnostic workup, time absent from work, and time to convalescence were longest, and the number of additional examinations was largest if only radiography had been performed and similar across the two MRI strategies. The mean total costs were 2231 Euro if only radiography was performed, 1820 Euro if additional MRI was performed in all cases, and 1697 Euro if MRI was performed only in the absence of a fracture on the radiograph.

Conclusion

Our results suggest that selective use of a short MRI examination following radiography in patients with acute knee injury who have no fracture on the radiograph saves costs and potentially increases effectiveness.

Introduction

Magnetic resonance imaging has proven to be a valuable imaging tool in the evaluation of traumatic knee injuries. The main reason for performing MRI after knee trauma is to assess if arthroscopy is indicated. This application of MRI has been studied extensively (1-4). MRI is, however, rarely used at first presentation of a patient with an acute knee trauma, although the information could be valuable, since physical examination shortly after a knee trauma is often hampered by swelling and pain. An important reason for not using MRI as an initial diagnostic tool in patients with acute knee injury is the generally high costs and long duration of the examination. MRI may, however, still play a significant role in the routine initial examination of recent knee injury. A requisite would be a relatively inexpensive MRI examination, of limited duration, which still provides sufficient information to influence the treatment strategy.
The costs and the duration of the MRI examination can be reduced if a low field dedicated extremity MRI system is used with a short scanning protocol. Although shortening of the scanning protocol would reduce image quality compared to a standard high field MRI examination, the information gained could be sufficient to reduce the time to completion of the diagnostic workup: the information could influence the decision whether follow-up of the patient is warranted, whether treatment is indicated, or whether the patient can be sent home without need of follow-up. If indeed a lesion would be detected, the early availability of this information could lead to earlier treatment of the patient and potentially to earlier recovery. This could lead to earlier resumption of work, resulting in a decrease of production losses, and thereby a decrease of costs to society.

The aim of this study was to assess the costs and effectiveness of performing a short MRI examination on a low field dedicated extremity MRI system in patients with an acute knee injury following radiography. We compared MRI in all patients and selective MRI in patients without a fracture on the radiograph to radiography only.

Methods

Study design

We performed a prospective pragmatic randomized controlled trial (RCT) in a university hospital, including patients with a traumatic knee injury, which had occurred within the preceding seven days before presentation. All patients were examined by an emergency physician, traumatologist or orthopaedic surgeon and radiography of the knee was ordered.

Exclusion criteria were: pre-existing complaints of the knee, compound fracture, substantial injury of head, back, thorax or abdomen, need for urgent treatment, and intoxication. The subjects were randomized across two diagnostic strategies consisting of plain radiography alone (strategy 1) and radiography followed by a short MRI examination (strategy 2). We modeled a third strategy consisting of plain radiography, followed by selective use of a short MRI examination, if no fracture was visible on the radiograph. The results of this strategy were estimated by using a composite of the results of patients from strategy 1 who showed a fracture on the radiograph and patients from strategy 2 who did not show a fracture on the radiograph. In this way the costs and effects of the third strategy could be obtained from the randomized patient groups.

Patients were randomized by drawing from consecutively numbered sealed envelopes containing computer generated random assignments. Block randomization was used with a block size of 20 to obtain equal numbers of patients in both strategies. Research staff and radiology technologists on service carried out the inclusion from 8.00 h to 23.00 h, 7 days a week. All patients were given written and oral information about the goal of the study, and all participating subjects gave informed consent. The results of the RCT are reported in accordance with the consort statement (5). The study was approved by the institutional review board and was performed in accordance with the Helsinki Declaration of 1964, amended in 1989.

Imaging technique and interpretation

All patients underwent anteroposterior and lateral radiography of the affected knee, and additional patellar or tunnel views if considered necessary. MRI was performed using a 0.2 T dedicated extremity MRI system (Artoscan M, Esaote S.p.a., Genoa, Italy). The scanning protocol is listed in table 1. The average scanning time was 6 minutes and total examination time, including MRI start-up and patient positioning was on average 15 minutes. The MRI was assessed by an experienced musculoskeletal radiologist (A.Z.G.) or by a resident on service during evenings and weekends. The result was reported to the treating physician immediately. If the MRI was assessed by a resident it was reassessed the next working day by A.Z.G. In case of a different interpretation the physician was informed.

Follow-up

Data were collected on utilization of medical resources, quality of life, and production losses caused by absence from work and time to convalescence. The follow-up period was 6 months.

Although 6 months is relatively short for outcome assessment, we expected that this period would be long enough for relevant differences across the strategies to emerge.

Table 1. Parameters of the short MRI protocol

	Plane	TR (ms)	TE (ms)	Matrix	Slice thickness (mm)	Acquisition time (min:s)
(Half Fourier) SE T1-w	Sagittal	490-720	18-24	192 x 216	5.0	0:49 - 1:08
TME PD and T2-w	Coronal	2150-2640	28-38/90	192 x 128	5.0	2:23 - 2:41
GE *	Coronal	500-595	15-16	192 x 128	5.0	0:48 - 0:53
STIR †	Sagittal	1060-1290	24-28	192 x 128	6.0/7.0	1:37 - 1:46

SE = spin echo, TME = turbo multi echo, PD = proton density, GE = gradient echo, STIR = short-tau inversion recovery, TR = repetition time, TE = echo time.
For all sequences: field of view : 200x140 mm, number of excitations :1
* Flip angle : 75 degrees
† Inversion time : 80-85 ms

The effect of the availability of more diagnostic information at the initial stage is not likely to cause a significant difference in costs beyond 6 months after the trauma. Questionnaires were sent to all subjects 1 week, 6 weeks, 3 months, and 6 months after inclusion. These questionnaires included quality of life measuring instruments as well as questions about utilization of medical resources outside our hospital, out-of-pocket expenses, days off work and time to convalescence. If questionnaires were not returned we interviewed the patient by telephone and urged them to return the quality of life questionnaires. Besides the information from the questionnaires, data were obtained from patient records and from the computerized hospital information system.

Measurement of effectiveness

Quality of life was measured using the EuroQol (6, 7) and the Short Form 36 Health Survey (SF-36) (8). Using the EuroQol it is possible to assign one preference based tariff, which can be calculated using a regression equation (9). The SF-36 consists of 36 questions covering eight domains (physical functioning, role-physical, bodily pain, general health, vitality, social functioning, role-emotional, and mental health). An algorithm is used to assign values to each domain.

From the moment a subject indicated on one of the four questionnaires that he or she did not have complaints of the injured knee anymore, we assumed that no further change in quality of life related to the initial trauma would occur and the EuroQol and SF-36 scores measured at that point in time were extrapolated to the remaining questionnaires. If a person indicated that he/she did not have complaints anymore but did not fill in the quality of life questions, mean values for the EuroQol and the eight SF-36 domains were used for extrapolation. These mean values were derived from questionnaires of all complaint-free subjects that filled in the health survey questionnaires. We used linear regression analysis to analyze if these mean values were influenced by age and sex and adjusted accordingly.
The time to completion of the diagnostic workup was defined as the time from initial presentation to the last diagnostic examination. The number of additional diagnostic procedures was assessed by reviewing the hospital information system and by information from the questionnaires or telephone inquiry in case the examinations had been performed outside our hospital. The number of days absent from work and the time to convalescence were obtained from the questionnaires or, in case the questionnaires were not returned, by telephone inquiry. A patient was regarded as convalesced if he or she did not have daily complaints anymore.

Measurement of costs

All costs relevant from the societal perspective were recorded with a follow-up period of 6 months. These costs included both medical and non-medical costs. Medical costs consisted of costs of diagnostic procedures and treatment both inside and outside the hospital as well as patient travel costs. Costs of initial treatment as well as diagnostic procedures were calculated using a bottom-up approach (10), taking into account the initial investment of equipment, additional costs during use, maintenance, years of use, discounting and annuitization (11), number of procedures per year, personnel costs, materials used, room rent, housekeeping, administration, and overhead costs. Costs were discounted at a rate of 3% per annum (12). Estimated actual costs of hospital visits and hospital admissions were obtained from the Dutch Council for Care Insurance (10). For costs of operations reimbursement tariffs as established by the Dutch Central Organ for Tariffs in Healthcare were used.
Out-of-pocket expenses and patient time costs were recorded as direct non-medical costs. Patient time was valued in monetary units (12) which was determined by multiplying the time spent on follow-up, diagnostic and therapeutic procedures (including travel time, and waiting time) by the average net income of subjects

stratified for age (source: Dutch Central Bureau for Statistics, data for both working and not working subjects in the Netherlands in the year 2000).
Costs associated with production losses were estimated using the 'friction cost method' as described by Koopmanschap et al and recommended in guidelines for cost-effectiveness analysis (12, 13). According to the friction costs method costs to society are only generated during the time it takes to replace a sick employee, the friction period. After this period no costs to society are generated, since the sick employee is replaced by someone drawn from the ranks of the unemployed. For short-term absence from work the loss of productivity may be low, because work may be performed by colleagues or work can be postponed and performed on return of the sick employee. For long term or permanent absence extra costs of maintaining production, costs of production loss, costs of filling the vacancy and costs of training the new employee are incorporated.
The costs of lost productivity per day absent from work were calculated using friction cost data estimated for working people in the Netherlands in 1998, based on sex and age (10), adjusted for the year 2000. The friction period for the year 1998 was estimated to be 4 months (10), but because of increased shortage on the labor market we estimated the friction period in 2000 to be 6 months (personal communication Koopmanschap). Costs were analyzed on an intention-to-diagnose-and-treat basis.
The effect of uncertainty of the involved parameters on the robustness of our results was studied in a sensitivity analysis. One-way sensitivity analysis was conducted by exploring a range from 50% to 200% for each parameter. Parameters that appeared sensitive in the one-way analysis were analyzed in a three-way sensitivity analysis.
The analyses were performed using Microsoft Excel 97 SR-2 (Microsoft Corp, Redmond, WA) and SPSS for Windows (release 10.0.0; SPSS Inc., Chicago, IL, USA) software packages. Because one of our strategies is a composite of the results of the other two, statistical testing for differences was not justified. We therefore chose to report 95% confidence intervals, but not p-values.

Results

From August 1999 to May 2001, 208 patients were included. A flow diagram of patients passing through the study is presented in figure 1. On review 18 patients did not fulfill all the inclusion criteria and one case was suspected of automutilation; these 19 patients were excluded from the analysis. Of the remaining 189 patients 96 patients were allocated to the MRI strategy and 93 patients to the reference strategy.

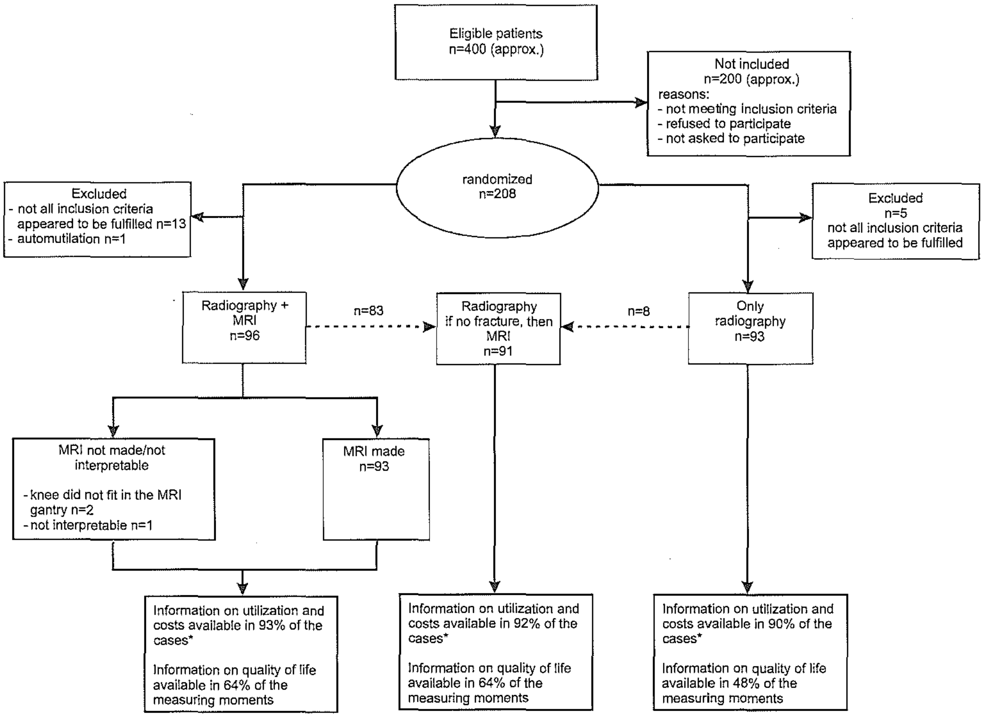


Figure 1. Flow diagram of subjects passing through the trial.

* the remaining percentage concerns patients that were not treated in our hospital, and information on possible treatment elsewhere is lacking.

Table 2. Baseline characteristics of the patient groups in the three strategies.

	Strategy 1 n=93	Strategy 2 n=96	Strategy 3 n=91
Age	34.7	32.2	33.7
Sex (male)	68 %	62 %	63 %
Fracture visible on X-ray	8	13	8
Type of trauma * (direct / indirect / unknown)	41 / 51 / 1	36 / 59 / 1	36 / 55 / 0

* numbers are the number of patients with the specified type of trauma
Strategy 1: only radiography
Strategy 2: radiography plus MRI
Strategy 3: radiography, if no fracture then MRI.

Table 3. Final diagnoses in 189 patients with recent knee injury.

	strategy 1	strategy 2	strategy 3
no abnormalities	0	2	2
contusion	18	24	25
distorsion*	21	14	14
fracture			
patella	4	6	4
tibiaplateau	0	2	0
fibula head avulsion	0	1	0
avulsion at the MCL origo	0	1	0
Segond fracture	1	0	1
osteochondral fracture	3	3	3
quadricepstendon rupture	1	0	0
medial meniscus tear	8	11	11
lateral meniscus tear	4	5	4
cruciate ligament rupture			
ACL partial	5	3	3
ACL total	7	14	12
PCL partial	0	1	1
PCL total	1	0	0
ACL+PCL total	1	0	0
ACL+PCL partial	1	0	0
collateral ligament rupture			
MCL partial	10	6	6
MCL total	3	4	3
LCL partial	0	1	1
LCL total	0	0	0
prepatellar bursitis	0	1	1
traumatized gonarthrosis	3	1	1
unclear	9	5	5

Strategy 1: only radiography
Strategy 2: radiography plus MRI
Strategy 3: radiography, if no fracture then MRI.
ACL=anterior cruciate ligament, PCL=posterior cruciate ligament,
MCL=medial collateral ligament, LCL=lateral collateral ligament.
* A distortion was defined as an indirect trauma (torsion, hyperextension, varus or valgus) without proven osseous, meniscal or ligamentous injury.

In 2 patients, allocated to the MRI strategy, the MRI was not made because the knee could not be positioned in the center of the magnet bore: one patient had a locked knee, and in one case the knee was too large to fit into the magnet. In one very obese patient the MRI was uninterpretable.

These three patients were analyzed in the strategy as originally assigned. The baseline characteristics are described in table 2. Eight patients in strategy 1 (only radiography) showed a fracture on the radiograph and 83 patients in strategy 2 (radiography plus MRI) showed no fracture on the radiograph. These 91 patients formed the subject group for strategy 3 (MRI if no fracture on radiography). The final diagnoses as assessed using all information including follow-up are listed in table 3.

Effectiveness

Linear regression analysis showed that the mean EuroQol and SF-36-domain scores of patients without complaints were not significantly influenced by age or sex, and therefore the mean values of the EuroQol and SF-36-domain values were used for imputation. The EuroQol tariffs are presented in figure 2.

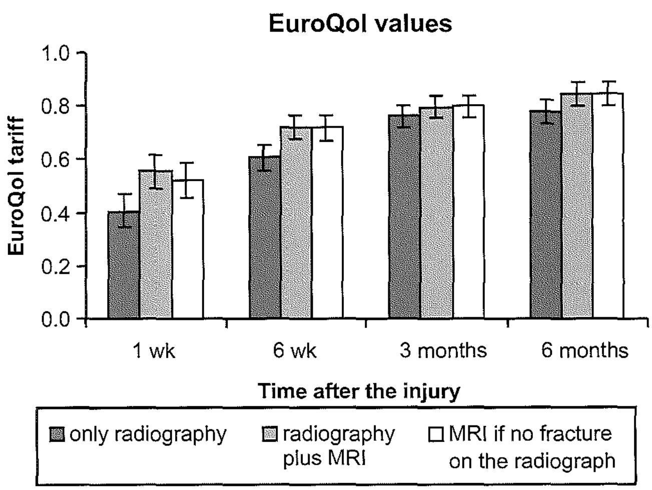


Figure 2. Mean EuroQol tariffs and 95% confidence intervals of the three strategies measured at 4 points in time after the initial injury.

Patients in strategy 2 (radiography plus MRI) showed a higher EuroQol tariff at one and six weeks after the injury (EuroQol tariff 0.56, 95% CI (confidence interval) 0.49-0.62 and 0.72, 95% CI 0.67-0.77 respectively) than patients in strategy 1 (only radiography; EuroQol tariff 0.41, 95% CI 0.33-0.48 and 0.61, 95% CI 0.54-0.67 respectively). In strategy 3 (MRI if no fracture visible on the radiograph) the EuroQol tariff at one week was slightly lower (EuroQol tariff 0.52, 95% CI 0.45-0.59) compared to strategy 2 (always MRI), whereas during follow-up the EuroQol tariffs of these two strategies were similar.

All the SF-36 domains, except for general health, demonstrated an increase in score over time (Figure 3).

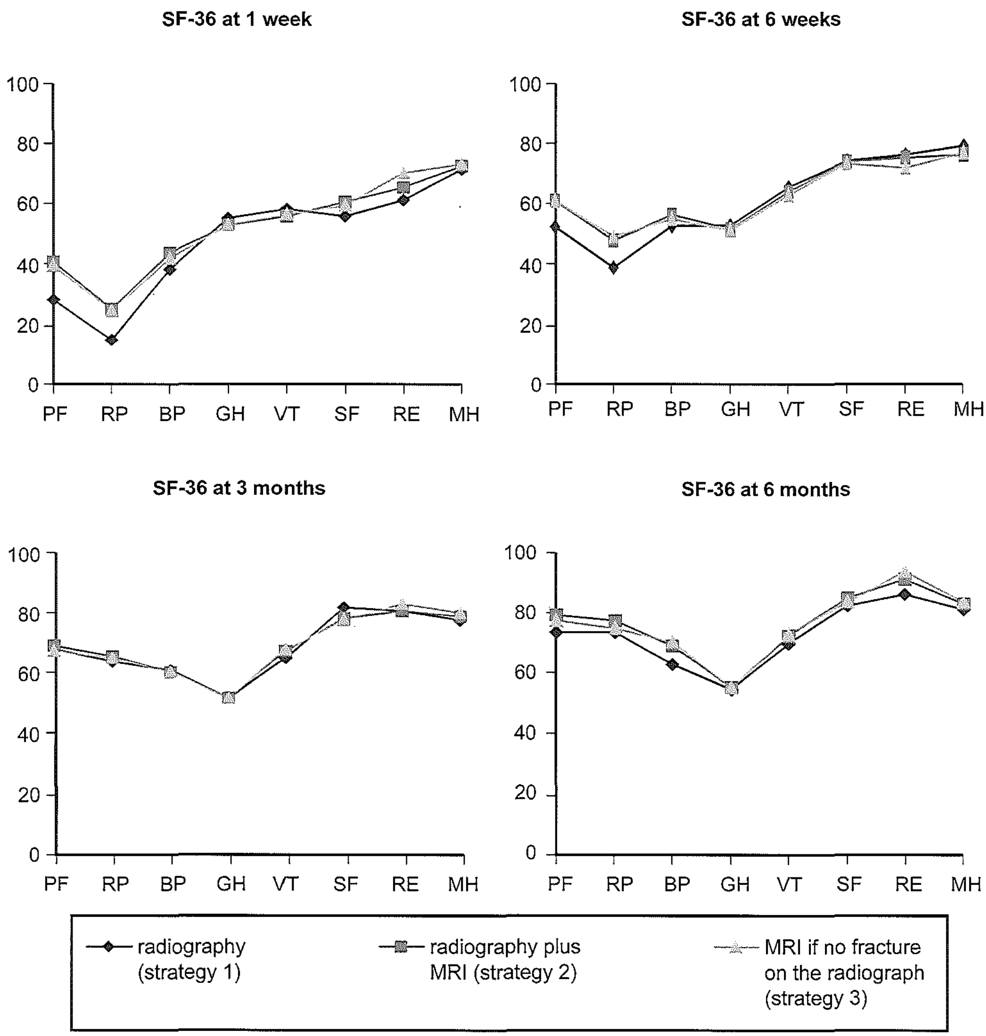


Figure 3. Mean SF-36 domain scores at 1week, 6 weeks, 3 months and 6 months after the injury.
PF= physical functioning, RP=role-physical, BP=bodily pain, GH=general health, VT=vitality, SF=social functioning, RE=role-emotional, MH=mental health.
Only the domains PF and RP at one and six weeks show a lower score in the 'radiography only' strategy compared to the other two strategies. Note: the lines between the points are only to improve visual comparability and do not have intrinsic meaning.

The two MRI strategies demonstrated similar SF-36 scores across all domains at all points in time. After one and six weeks physical functioning and physical role functioning were higher in the MRI strategies compared to radiography alone. However, only the difference in physical functioning after one week was statistically significant. The scores of all the other domains were very similar across the three strategies at all points in time.The time to completion of the diagnostic workup, the duration of absence from work, and the time to convalescence were shortest in strategy 3 (Table 4). The number of additional procedures was almost the same in the two strategies with MRI, which was considerably shorter than in the strategy with only radiography.

Costs

Medical costs were highest in the strategy with MRI in all patients (strategy 2) and lowest if no MRI was performed at all (strategy 1) (Table 5). Travel costs were relatively low and similar across the three strategies. The friction costs constituted the largest cost factor in all strategies.

Table 4. Effectiveness results.

	Strategy 1 (95% CI) n=93	Strategy 2 (95% CI) n=96	Strategy 3 (95% CI) n=91
Mean time to last diagnostic procedure (days)	17.3 (9.3-25.2)	3.5 (0.0-7.6)	2.0 (0.0-4.1)
Number of additional diagnostic procedures (absolute numbers)	35	9	10
Mean duration of absence from work (days)	12.4 (7.5-17.3)	9.6 (5.9-13.2)	8.6 (5.0-12.3)
Mean time to convalescence (days)	76.6 (54.6-98.5)	66.2 (50.2-82.3)	60.4 (44.8-76.0)

Strategy 1: only radiography
Strategy 2: radiography plus MRI
Strategy 3: radiography, if no fracture then MRI
95% CI = 95% confidence interval

Total costs were lowest if MRI was used selectively in patients without a fracture on the radiograph (strategy 3), which was mainly caused by a reduction in friction costs. Total costs were highest in the strategy with only radiography, which again was mainly attributable to high friction costs.

Table 5. Mean costs associated with the initial knee injury.

	Strategy 1 Euro (95% CI)	Strategy 2 Euro (95% CI)	Strategy 3 Euro (95% CI)
Medical costs	606 (462-751)	691 (486-897)	640 (520-761)
Travel costs	16 (14-19)	17 (15-20)	18 (15-20)
Friction costs	1510 (896-2123)	993 (556-1430)	944 (503-1385)
Time costs	99 (58-140)	118 (29-208)	94 (57-132)
Total costs	2231 (1561-2901)	1820 (1280-2360)	1697 (1205-2188)

Strategy 1: only radiography
Strategy 2: radiography plus MRI
Strategy 3: radiography, if no fracture then MRI
95% CI = 95% confidence interval

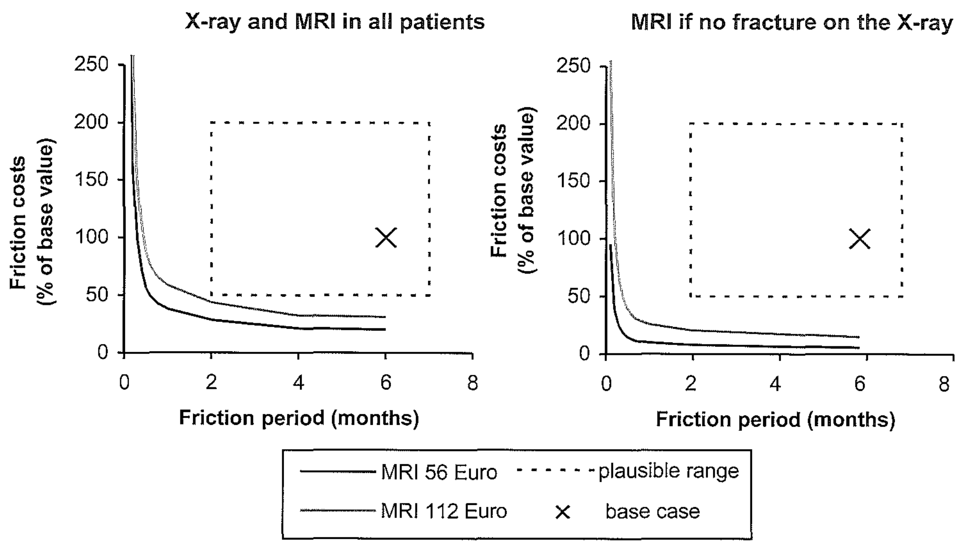


Figure 4. Three-way sensitivity analysis on friction costs and friction period for the base value of the short MRI examination (56 Euro) and twice this base value. Radiography and MRI in all patients compared with radiography only (a); radiography in all patients followed by MRI only if no fracture is visible on the radiograph compared with radiography only (b). In the area above the threshold lines the strategy that includes MRI generates less costs than the reference strategy, in the area below the threshold lines the strategy including MRI generates more costs than the reference strategy. A plausible range of friction costs and friction period is indicated.

Table 6. One-way sensitivity analysis of the mean total cost difference between the intervention and the reference strategy.
The cost difference between the two strategies was recalculated for each single variable at 50% and 200% of the original value of that variable, to assess the influence of variation of that variable on the total cost difference. Costs of the strategy that did not include MRI were always highest. The baseline cost difference is the cost difference between the strategies with all variables fixed at 100%. The range is a measure of sensitivity of the cost difference to that variable. Only the most sensitive variables are presented. Costs are in Euro.

	Cost difference between strategy 1 and strategy 2			Cost difference between strategy 1 and strategy 3		
Variable	50 %	200 %	range	50 %	200 %	range
baseline (all variables 100%)		412			534	
friction costs	153	929	776	252	1099	847
friction period	264	412	148	372	534	162
short MRI examination	440	355	85	560	483	77
short MRI examination	440	355	85	560	483	77
outpatient visit	425	385	40	546	511	35
time costs	421	392	29	532	539	7
hospital day	418	398	20	517	569	52
physiotherapy	416	402	14	545	514	31
standard MRI examination	408	418	10	525	553	28
operative therapy	410	415	5	531	541	10
radiography	413	408	5	535	533	2

Strategy 1: only radiography
Strategy 2: radiography plus MRI
Strategy 3: radiography, if no fracture then MRI.

Sensitivity analysis

In a one-way sensitivity analysis the cost difference between the strategies including MRI (strategy 2 and 3) and the strategy with only radiography (strategy 1) was sensitive to the estimated friction costs per friction period, to the friction period itself, and to the costs of the short MRI examination (Table 6). Three-way sensitivity analysis of friction costs, friction period, and costs of the short MRI examination (Figure 4) demonstrated that within plausible ranges of these variables the strategies that included MRI generated less costs than the strategy with only radiography. This conclusion was even more robust for strategy 3 (MRI if no fracture on radiograph) (Figure 4b) than for strategy 2 (always MRI) (Figure 4a).

Discussion

MRI is seldomly used as an initial investigation tool in acute knee injury because of the high costs of the examination. We studied the costs and effectiveness of applying a limited MRI examination in patients with an acute knee injury. Two MRI strategies were analyzed, namely performing MRI in all patients and selective use of MRI if no fracture is demonstrated on the radiograph. The rationale behind the latter strategy is that we considered it plausible that the addition of MRI to radiography in patients who already show a fracture on the radiograph has limited value for the (initial) treatment: in most cases treatment will be determined by the fracture. The strategies were compared to the strategy of radiography alone.

Performing this study we assumed that the early diagnostic information would result in more adequate and earlier treatment. Although this is true for many cases, we recognize that there is a potential danger of overtreatment. Some meniscal tears may heal or become symptomless without surgery (14). Still, if the short MRI examination leads to an overall reduction in costs without a reduction in quality of life, the benefits of more accurate early diagnostic information outweighs the risk of overtreatment.

The application of MRI in patients without a fracture on the radiograph resulted in a slight improvement in quality of life during the first 6 weeks compared to using radiography only. The quality of life was almost identical for the two MRI strategies. Time to diagnosis, duration of absence from work, and time to convalescence were all slightly shorter compared to MRI in all patients and substantially shorter compared to only radiography.

Medical costs were lowest if only radiography was performed. As expected, if an MRI was performed selectively in the absence of a fracture, medical costs were lower than if MRI was performed in all patients. Thus, the application of a short MRI examination did not lead to a reduction in costs of subsequent diagnostic and therapeutic procedures. The main reduction in costs was brought about by a reduction in lost productivity, which was similar for both MRI strategies. The difference in time costs was small across strategies because the mean time spent for diagnostic and therapeutic procedures was similar. Travel costs were similar and low for all groups. This is plausible in the Dutch situation, since the density of

hospitals in the Netherlands is high, and travel distance to the nearest hospital is generally short.
Our conclusions were robust in sensitivity analysis. A three-way analysis of the most sensitive parameters (friction costs, friction period, and costs of the short MRI examination) did not influence our conclusions. Our results were, however, influenced by heterogeneity of the patient populations, expressed by the wide confidence intervals of the costs (Table 5). This heterogeneity within the patient groups was to be expected since we considered all patients with traumatic knee injury, ranging from a mild injury, without need for treatment to trauma with extensive internal derangement. Especially friction costs were subject to heterogeneity in the studied population. Since a short period off work may generate considerable costs, absenteeism caused high mean friction costs in spite of the fact that only a small number of patients was absent from work.

A limitation of the study is the fact that although inclusion was intended to be consecutive, about half of the eligible patients was not included. This was caused by patients who refused to participate and by patients that were not asked to participate; because inclusion took place from 8.00 - 23.00 h, seven days a week, many different radiology technologists on service were involved in the recruitment, which was forgotten regularly.

In many studies the response rate to mailed questionnaires is about 60 - 65% (15-22); in our study the response rate of patients that underwent both radiography and MRI was within that range, but the response rate of patients that only underwent radiography was lower, which may have biased the results on quality of life. For data on costs we were much less dependent on the questionnaires, since we could obtain most data from the patient records and computerized hospital information system, sometimes complemented by telephone inquiry.
The diagnostic strategy with selective use of MRI if no fracture was visible on the radiograph was modeled, since this group was not incorporated in the trial. Modeling created limitations in the comparison of the strategies, since differences between the modeled and the observed strategies could not be tested statistically. However, we chose to use a modeled strategy because it created the opportunity to analyze a realistic strategy beyond the scope of the original trial. The results indicate that the modeled strategy, MRI in patients without a fracture on the radiograph, may be more cost-effective than MRI in all patients. Because the results were obtained from a post hoc modeling analysis, they should ideally be evaluated in a prospective trial.

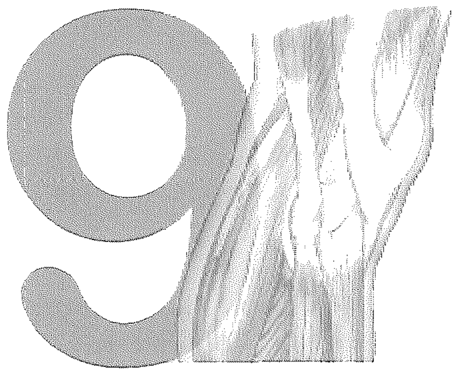
Conclusion

Our results suggest that selective use of a short MRI examination following radiography in patients with acute knee injury who have no fracture on the radiograph saves costs and potentially increases effectiveness.

References

1. Oei EH, Nikken JJ, Verstijnen AC, Ginai AZ, Myriam Hunink MG. MR imaging of the menisci and cruciate ligaments: a systematic review. Radiology 2003;226(3):837-48.
2. Elvenes J, Jerome CP, Reikeras O, Johansen O. Magnetic resonance imaging as a screening procedure to avoid arthroscopy for meniscal tears. Arch Orthop Trauma Surg 2000;120(1-2):14-6.
3. Munk B, Madsen F, Lundorf E, Staunstrup H, Schmidt SA, Bolvig L, et al. Clinical magnetic resonance imaging and arthroscopic findings in knees: a comparative prospective study of meniscus anterior cruciate ligament and cartilage lesions. Arthroscopy 1998;14(2):171-5.
4. Weinstabl R, Muellner T, Vecsei V, Kainberger F, Kramer M. Economic considerations for the diagnosis and therapy of meniscal lesions: can magnetic resonance imaging help reduce the expense? World J Surg 1997;21(4):363-8.
5. Moher D, Schulz KF, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. Jama 2001;285(15):1987-91.
6. EuroQol--a new facility for the measurement of health-related quality of life. The EuroQol Group. Health Policy 1990;16(3):199-208.
7. Brooks R. EuroQol: the current state of play. Health Policy 1996;37(1):53-72.
8. Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992;30(6):473-83.
9. Dolan P. Modeling valuations for EuroQol health states. Med Care 1997;35(11):1095-108.
10. Oostenbrink JB, Koopmanschap MA, Rutten FFH. Handleiding voor kostenonderzoek, methoden en richtlijnprijzen voor economische evaluaties in de gezondheidszorg. Amstelveen: College voor zorgverzekeringen; 2000.
11. Hunink MGM, Glasziou PP. Decision making in health and medicine - integrating evidence and values. Cambrige: Cambridge University Press; 2001.
12. Gold MR, Siegel JE, Russell LB, Weinstein MC. Cost-effectiveness in health and medicine. Oxford: Oxford University Press; 1996.
13. Koopmanschap MA, Rutten FF, van Ineveld BM, van Roijen L. The friction cost method for measuring indirect costs of disease. J Health Econ 1995;14(2):171-89.
14. Solomon DH, Simel DL, Bates DW, Katz JN, Schaffer JL. The rational clinical examination. Does this patient have a torn meniscus or ligament of the knee? Value of the physical examination. Jama 2001;286(13):1610-20.

15. Kotaniemi JT, Hassi J, Kataja M, Jonsson E, Laitinen LA, Sovijarvi AR, et al. Does non-responder bias have a significant effect on the results in a postal questionnaire study? Eur J Epidemiol 2001;17(9):809-17.
16. Becker A, Seitz R, Jacobi E, Leidl R. [Cost evaluation by a patient questionnaire: pilot study of a weekly cost diary] Kostenmessung durch Patientenbefragung: Pilotstudie zu einem Kostenwochenbuch. Rehabilitation (Stuttg) 2001;40(1):12-20.
17. Biering-Sorensen F, Biering-Sorensen M. Sleep disturbances in the spinal cord injured: an epidemiological questionnaire investigation, including a normal population. Spinal Cord 2001;39(10):505-13.
18. Heath AC, Howells W, Kirk KM, Madden PA, Bucholz KK, Nelson EC, et al. Predictors of non-response to a questionnaire survey of a volunteer twin panel: findings from the Australian 1989 twin cohort. Twin Res 2001;4(2):73-80.
19. Bakker E, Van Gool JD, Van Sprundel M, Van Der Auwera C, Wyndaele JJ. Results of a questionnaire evaluating the effects of different methods of toilet training on achieving bladder control. BJU Int 2002;90(4):456-61.
20. Campbell MJ, Waters WE. Does anonymity increase response rate in postal questionnaire surveys about sensitive subjects? A randomised trial. J Epidemiol Community Health 1990;44(1):75-6.
21. Lund E, Gram IT. Response rate according to title and length of questionnaire. Scand J Soc Med 1998;26(2):154-60.
22. Maconochie N, Doyle P, Davies G, Lewis S, Prior S, Sampson P, et al. The study of reproductive outcome and the health of offspring of UK veterans of the Gulf war: methods and description of the study population. BMC Public Health 2003;3(1):4.



Value of Information Analysis Directing Further Research: Assessing the Use of MRI for Acute Knee Trauma

Bas Groot Koerkamp, MD MSc[1], Jeroen J. Nikken, MD[2, 1],
Theo Stijnen, PhD[1], Abida Z. Ginai, MD, PhD[2],
M.G. Myriam Hunink, MD, PhD[1,2]

From the Program for the Assessment of Radiological Technology (ART Program).
Department of Epidemiology and Biostatistics[1], Department of Radiology[2],
Erasmus MC, University Medical Center, Rotterdam, The Netherlands.

Abstract

Purpose

The purpose of this study was to illustrate the usefulness of value of information analysis to direct further research concerning the use of MRI for acute knee trauma.

Methods

We performed value of information analyses in a probabilistic model comparing no MRI versus MRI versus selective MRI for acute knee trauma from the societal perspective, using data from a randomized controlled trial.

Results

The total expected value of perfect information (EVPI) of all parameters was 83 euro per patient, resulting in a population EVPI of 7.6 million euro. We found that only four parameters (out of 108) together were responsible for 83% of the benefit to gain with more research. These parameters represent the number of days off work due to the knee injury for each of the three strategies and the number of hospital days for the intervention strategy.

Conclusions

This application illustrates the usefulness and practicality of value of information analysis to guide further research. Instead of performing another trial to gather better estimates of all 108 parameters, our results suggest that the next study should focus on obtaining better estimates for the number of days off work for each strategy.

Introduction

The main objective of a randomized controlled trial (RCT) or medical decision analysis is to identify the optimal treatment or diagnostic work-up. If an RCT finds a p-value of less than the arbitrary 0.05, investigators tend to conclude that the issue is settled. For higher p-values it is almost invariably concluded that more research is needed. However, the expected benefit of more research is not simply high for a p-value exceeding 0.05 and small for a p-value below 0.05. The justification of more research depends on such diverse factors as the number of future patients that could benefit, the costs of future research, and the probability and consequences of a false claim. Apparently the results of an RCT raise a second question: should we perform more research? (1-4)

More research seems only justified if the expected benefit of more research for future patients exceeds the expected costs of the research itself. Value of information analysis is an analytical tool that can estimate the expected value of more research and therefore direct further research. If the benefit of more research is not explicitly assessed, a non-significant trial result should conclude that there is insufficient evidence to identify the best treatment, instead of the universal "more research is needed".

In a randomized controlled trial we assessed the use of MRI for acute knee trauma. After completion of this study we were left with the decision whether further research should be performed to decrease uncertainty about the optimal decision. The purpose of this study was to illustrate the usefulness of value of information analysis to direct further research concerning the use of MRI for knee trauma.

Methods

Randomized controlled trial

In a diagnostic randomized controlled trial we enrolled consecutive patients between August 1999 and May 2001 with recent knee injury referred to the radiology department for imaging. Patients were randomized across two diagnostic strategies consisting of plain radiography alone (the reference group) and radiography followed by a short MRI examination (the intervention group). A third strategy was modeled, including those randomized patients in the reference group with a fracture and those in the intervention group without a fracture. The modeled strategy thus represented patients that underwent an MRI examination only if the plain radiography was negative.

During the 6 months of follow-up, quality of life was measured with the EuroQol. All costs relevant from the societal perspective were recorded during the follow-up period. These costs included both medical and non-medical costs.

Medical costs consisted of costs of diagnostic procedures and treatment both inside and outside the hospital as well as patient travel costs.
Costs of initial treatment as well as diagnostic procedures were calculated using a bottom-up approach (5), taking into account the initial investment of equipment, additional costs during use, maintenance, years of use, discounting and annuitization (6), number of procedures per year, personnel costs, materials used, room rent, housekeeping, administration, and overhead costs. Costs were discounted at a rate of 3% per annum (7). Estimated actual costs of hospital visits and hospital admissions were obtained from the Dutch Council for Care Insurance (5). For costs of operations reimbursement tariffs as established by the Dutch Central Organ for Tariffs in Healthcare were used.
Out-of-pocket expenses and patient time costs were recorded as direct non-medical costs. Patient time was valued in monetary units (7) which was calculated by multiplying the time spent on follow-up, diagnostic and therapeutic procedures (including travel time, and waiting time) by the average net income over all randomized subjects (source: Dutch Central Bureau for Statistics, age-specific data for both working and not working subjects in the Netherlands in the year 2000).
Costs associated with production losses were estimated using the 'friction cost method' as recommended in guidelines for cost-effectiveness analysis (7,8). According to the friction costs method, costs to society are only generated during the time it takes to replace a sick employee, the friction period. After this period no costs to society are generated, since the sick employee is replaced by someone drawn from the ranks of the unemployed. For short-term absence from work the loss of productivity may be low, because work may be performed by colleagues or work can be postponed and performed on return of the sick employee. For long term or permanent absence extra costs of maintaining production, costs of production loss, costs of filling the vacancy and costs of training the new employee are incorporated.
The costs of lost productivity per day absent from work were calculated using friction cost data estimated for working people in the Netherlands in 1998, based on sex and age (5), adjusted for the year 2000. The friction period for the year 1998 was estimated to be 4 months (5), but because of increased shortage on the labor market we estimated the friction period in 2000 to be 6 months (personal communication Koopmanschap). The friction cost was determined by multiplying the number of days absent from work by the average costs, over all randomized subjects, of lost productivity per day absent from work.

Probabilistic model

We developed a probabilistic model using the data from our RCT(9, 10). A probabilistic model requires specifying distributions for all estimated model parameters in order to represent their uncertainty. Quality of life was not considered in the model since a difference was not anticipated. All medical costs, as well as travel and friction costs, were split up in unit costs (e.g., the costs of a cast or a day inpatient stay) and the mean number of units per patient (resource use) for each strategy (e.g., 5 casts for 100 patients or 6 inpatient days for 100 patients). The unit costs are independent of the strategy, contrary to the mean number of units per patient. Patient time costs were split up in unit time (e.g., time needed for an outpatient visit), mean number of units (e.g., mean number of outpatient visits per patient) and the time costs per minute.

Since unit costs are specified and not estimated parameters, we treated them as fixed rather than stochastic(11). Distributions representing the uncertainty about the mean number of units per patient for each unit were derived from patient level data. All distributions in the model represent the uncertainty about the estimated mean value of the parameters (standard error of the mean), not of the distribution in the sampled population (standard deviation). The distribution of resource use in a sampled population is usually very skewed, which can be described, for instance, with a Gamma distribution. The uncertainty about the mean values in the population, however, was represented with normal distributions since the central limit theorem held.

In the RCT, data on resource use were collected for 36 parameters for both strategies. The modeled strategy added another 36 parameters, resulting in a total of 108 resource parameters. For each of the 36 parameters the unit cost was specified, being the same for each strategy. For some of the 36 parameters (e.g., an outpatient visit) the unit time was specified. Twenty parameters represented rarely used resources (e.g., ultrasound of the knee or surgery for a quadriceps rupture) and were therefore aggregated into 4 groups. Instead of estimating the mean number of resources per parameter, the mean aggregated costs were estimated in these groups. This resulted in a total of 20 parameters per strategy (36-20+4) in the model.

The simultaneous uncertainty in all parameters was represented by a multivariate normal distribution with variances equal to the estimated squared standard errors of the mean and correlations equal to the estimated correlations between the different parameters. We assumed that the correlations between the parameters were identical for all strategies. Therefore, one correlation matrix was calculated using all randomized patients.

We performed second-order Monte Carlo simulations (1,000,000 iterations) to propagate uncertainty represented by the distributions through the model, resulting in distributions for the total costs of each strategy. These distributions should have the same mean and standard errors as the classical statistical analysis of patient data (i.e., calculating mean and standard error of total costs per patient). We then calculated the probability that the mean optimal strategy was optimal.

Value of information analysis

We performed value of information analysis as described by others in recent publications(1, 12-14). Total expected value of perfect information (total EVPI) was first calculated. Total EVPI can be interpreted as the expected benefit of an infinitely large RCT, reducing uncertainty to zero. Value of information is always expressed in the same units as the outcome of the strategies (i.c. euro).

We then estimated the annual number of patients expected to benefit from more information, the effective life time (T) of the short MRI examination, and the discount rate, in order to calculate the population EVPI(3):

$$\text{EVPI} * \sum_{t=1}^{T} \frac{\text{Annualpopulation}}{(1+\text{discountrate})^{t}}$$

If the population EVPI is smaller than the estimated research costs, more research is not justified.

Next, by calculating partial expected value of perfect information (partial EVPI) we estimated the contribution of one parameter, or sets of parameters, to overall uncertainty. The benefit of knowing the parameters of interest without uncertainty was calculated. Although the model is a linear function of the parameters, a two-level Monte Carlo simulation was necessary since the parameter(s) of interest and the remaining parameters were correlated.

Instead of sampling all parameters in a future study, key parameters were identified that were responsible for most of the decision uncertainty. These key parameters could be targeted specifically in either an RCT or an observational study.

All analyses were performed in S-PLUS 6.0 (Professional Release 1, Insightful Corp.).

Results

Randomized controlled trial

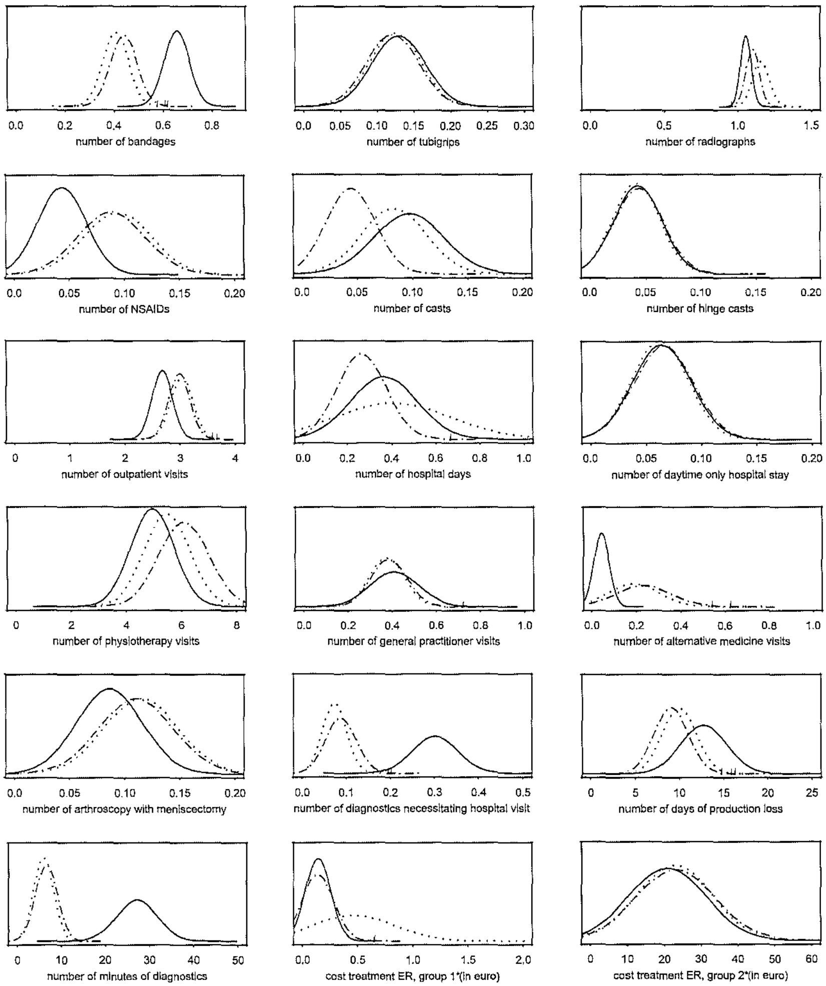
189 Patients were randomized, of whom 93 were allocated to the radiograph only strategy and 96 underwent an additional MRI of the knee. 91 Subjects represented the third (modeled) strategy. The mean total costs were 1944, 1774 and 1621 euro per patient, respectively. These means did not differ statistically significantly. Friction and medical costs were responsible for most of the total costs. Time and travel cost together were responsible for less than 10% of the total costs.

Model parameters

Table 1 shows a list of the values of the specified parameters (i.e., unit cost and unit time). The strategy-specific distributions of 20 estimated parameters, indicating mean number of units, including four aggregated parameters indicating mean costs, are presented by probability density plots in Figure 1.

Table 1. Values (i.e., unit costs and unit times) of specified parameters

Description	Value
Unit costs (in euro)	
bandage	5
tubigrip	2
NSAID	10
cast	38
hinge cast	57
outpatient visit	74
hospital day (24 hours)	340
daytime only hospital day	227
short MRI examination	56
radiograph (including radiologist interpretation)	15
physiotherapy visit	18
general practitioner visit	17
alternative medicine visit	18
arthroscopy and meniscectomy	480
travel cost emergency room	3
travel cost hospital	3
travel cost other	1
friction cost per day	96
time cost per minute	0.09
Unit time (in minutes)	
emergency room visit	90
short MRI examination	30
outpatient visit	60
general practitioner visit	45
alternative medicine/physiotherapy visit	60
hospital day	1440
daytime only hospital day	480



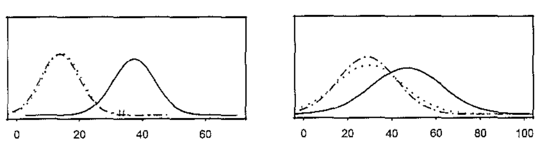


Figure 1. Probability density plot, representing uncertainty about the mean value of the 20 estimated parameters for each strategy. The solid line represents the reference strategy, the dotted the intervention strategy and the dashed line the modeled strategy. The goups refer to aggregated parameters of rarely used resources:

* Clustered variables:
group 1: punction, Fentanyl, stitching, joint drainage
group 2: brace (two types), upper-leg splint, upper-leg cast, upper-leg softcast
group 3: high-field MRI, computer tomography, ultrasound, exam in general anesthesia, arthroscopy
group 4: surgery: osteochondritis dissecans, arthroscopy without meniscectomy, quadriceps rupture, collateral ligament rupture, patella fracture, tibia plateau fracture, cruciate ligament rupture

These plots give an impression of the uncertainty around the mean values of the sampled parameters. The reference strategy (radiograph only) had a significantly higher mean value for parameters associated with additional diagnostic tests (i.e., costs of other diagnostic tests, number of minutes of diagnostic tests and number of diagnostic tests necessitating hospital visit). Since these patients did not undergo an MRI at their emergency room visit, they were more likely to need additional diagnostic work-up later during follow-up. The number of bandages is the only other parameter differing significantly between strategies, for which we do not have an explanation other than chance.

Probabilistic sensitivity analysis

Figure 2, a probability density plot, illustrates the uncertainty around the mean of the total costs for each strategy. For each strategy two curves are plotted: the wider distributions take the correlations between the parameters into account. The costs of surgical treatment, for example, are highly positively correlated with the number of hospital days. If correlations were not modeled, the overall decision uncertainty was underestimated. Taking correlations into account, the model found exactly the same mean and confidence interval as the classical statistical analysis of the RCT data.

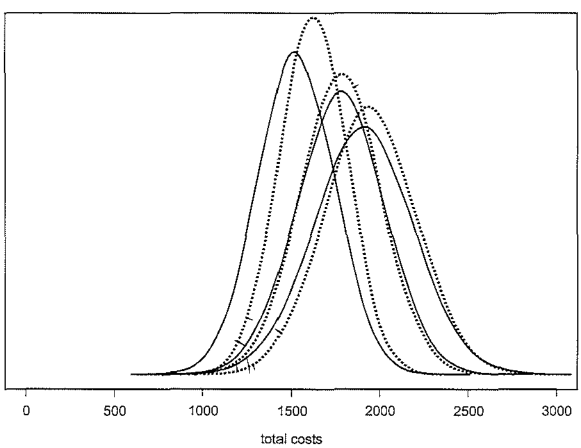


Figure 2. Probability density plot. The solid and wider distributions take correlations between parameters into account, the dotted distributions do not. From left to right: distribution of the modeled strategy, the intervention strategy and the reference strategy.

The modeled strategy had the lowest mean total cost and was the optimal strategy in 61% of the simulations. Therefore, if based on this RCT the modeled strategy would be selected, there would be a probability of 39% that the wrong (i.e., more costly) strategy is introduced.

Total value of perfect information

The total expected value of perfect information was 83 euro. This implies that an infinitely large future RCT is expected to reduce the mean total costs of a knee-injured patient with 83 euro on the current mean of 1621 euro per patient. The distribution of the total value of information was highly skewed as shown in the histogram in Figure 3. The first bar of the histogram is so high since in the 61% of the simulations in which the modeled strategy was optimal, the value of information was zero.

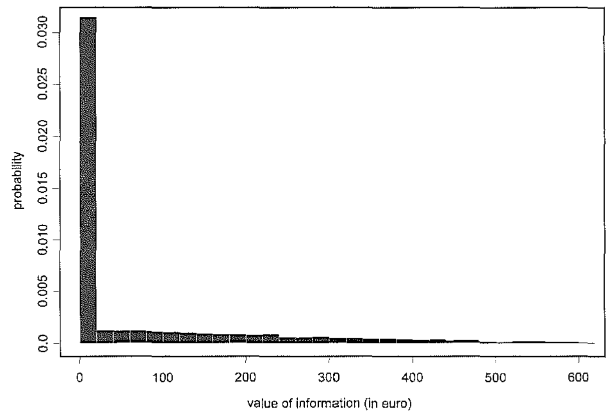


Figure 3. Histogram, scaled as a probability density function of the total value of information. The mean, or expected, value of information is 83 euro.

Figure 4 presents the "reversed" cumulative density curve. The y-axis shows the probability that the value of perfect information exceeds the value on the x-axis. Since the optimal strategy was optimal in 61% of the analyses, the probability that the realized value of perfect information will be larger than zero was only 39%. The probability that the value of perfect information will exceed the expected value of perfect information of 83 euro was about 29%.

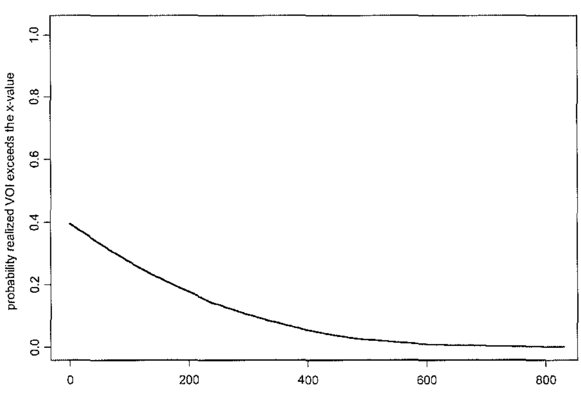
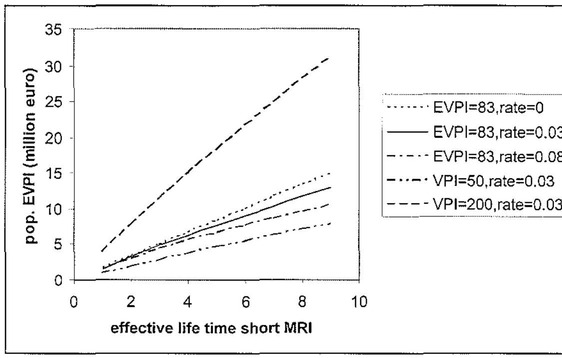


Figure 4. "Reversed" cumulative density curve. The y-axis shows the probability that the value of information exceeds the value on the x-axis.

Population EVPI

We estimated the number of future patients expected to benefit from more research at 20,000 per year for the Netherlands. There is considerable uncertainty about the effective lifetime of the short MRI examination as used in the intervention strategy. In Figure 5 the population EVPI is calculated over a plausible range for the effective lifetime. This was repeated for different values of the discount rate and the value of information for one patient. For an effective life time of 5 years and a discount rate of 3%, the population EVPI was 7.6 million euro.

Figure 5. Population EVPI as a function of the effective life time of the short MRI examination. The effects of different discount rates and values of perfect information (VPI) are explored.



Partial EVPI

Figure 6 presents the partial EVPI per patient for selected individual parameters and sets of parameters.

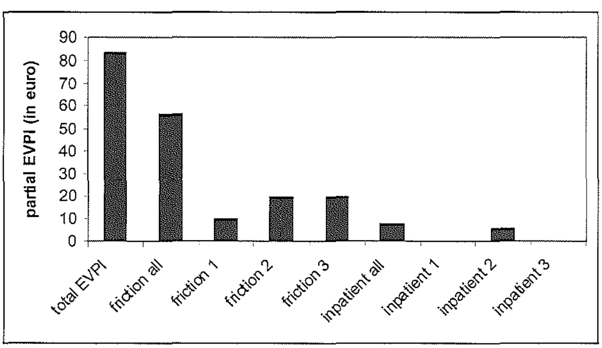


Figure 6. Partial EVPI per patient of individual parameters and sets of parameters. "Friction" refers to the number of days of production loss. "All" refers to all strategies simultaneously. 1,2 and 3 refer to the reference strategy, the intervention strategy and the modeled strategy, respectively. "Inpatient" refers to the number of hospital days.

Of the 108 (36 per strategy) parameters only 4 parameters were responsible for 83% of the total EVPI. The parameters representing the number of days of production loss for each strategy were responsible for most of the expected benefit to gain with more research. Reducing the uncertainty around the mean number of hospital days for the intervention strategy is also expected to be beneficial. The partial EVPI for the mean number of hospital days for the other strategies was close to zero. Note that figure 1 shows that the mean number of hospital days had the highest a priori uncertainty for the intervention strategy.
The partial EVPI of two parameters was not the same as the sum of the partial EVPI of each parameter separately. For example, the partial EVPI for the mean number of hospital days was 8.1 euro for the intervention and the third strategy together, but for the intervention strategy alone 6.1 euro and for the third strategy alone 0.02 euro.

Discussion

Patients undergoing an MRI in the emergency room setting for acute knee trauma, conditional on a negative radiograph (i.e., no fracture), incurred the least costs from a societal perspective during the 6 months of follow-up (but not significantly less). The total expected value of perfect information was 83 euro per patient, resulting in an estimated population EVPI of 7.6 million euro for the Netherlands. This population EVPI was sufficiently high to justify exploration of partial EVPI. We identified 4 key parameters that were together responsible for 83% of the total expected benefit to gain with more research. These 4 parameters were the mean number of days of production loss for each strategy and the mean number of hospital days for the intervention strategy. This implies that if we could get perfectly accurate estimates of these four parameters in a research project that would cost less than 83% of 7.6 million euro, the research project is expected to be cost-effective. Perfect estimates can only be derived with an infinitely large sample, however, and therefore we should assess the so-called expected value of sample information (EVSI) to estimate the expected value of different sample sizes(13).
Both in RCTs and decision models, not only effectiveness, but also costs and uncertainty are increasingly assessed with so-called stochastic cost-effectiveness analyses. Economic evaluations can help us decide between the many innovations, forced by the limited health care budget. However, the objective of assessing uncertainty is rarely stated explicitly.
The principle of decision theory states that decisions ought to be based on expected utilities (i.e., the mean outcomes)(15, 16). Claxton articulated this principle by stating that uncertainty around the expected outcome measure is irrelevant in

choosing the optimal strategy(1, 3): no matter what the p-value is, the strategy with the best mean outcome should be implemented. Clinical trials are not performed to draw inference, but to make decisions. Accordingly, assessing uncertainty can only be of value in deciding whether further research is necessary. The latter decision may be based on sensitivity analyses or acceptability curves, which both quantify the confidence we have that the optimal strategy really is the best and therefore whether more research is needed. But even the sophisticated probabilistic sensitivity analysis (PSA) can give an incorrect impression of the importance of the decision uncertainty: a high probability that another strategy than the optimal strategy is best, does not justify more research if the other strategy is not likely to be substantially better. PSA only assesses the probability of a wrong claim, not the consequences in costs or health benefit. For some time, value of information (VOI) analysis has been proposed to explicitly assess the value of performing more research regarding medical decisions. Nevertheless the vast majority of decision analyses explore uncertainty for ambiguous reasons instead of performing a VOI analysis to direct further research.

Ideally, a VOI analysis is performed in a decision model representing all current knowledge and uncertainty. We only used data from our RCT and therefore might have overestimated the expected value of information. Synthesizing and extrapolating all available evidence in a decision model, however, has the disadvantage of inevitable speculation and judgment.

One important judgment call we made concerns the validity of the modeled strategy. Though this strategy was not an arm of the original RCT, we believe it is unbiased. Another assumption is that beyond the follow-up period of 6 months we do not expect a difference in costs between strategies.

We also assumed that the unit costs of an MRI, the time costs per minute and the friction costs per day are specified and therefore fixed. However, the unit cost of an MRI, for example, was calculated with a bottom-up approach. Subunits (e.g., equipment use) were sampled, yielding a different MRI cost for each patient. The distribution of this subunit use could be more accurately estimated in a larger study. We did not take this additional uncertainty into account, possibly slightly underestimating the total EVPI.

An influential parameter in a VOI analysis is the number of patients expected to benefit. Figure 5 illustrates how an increase in effective life time results in a steep increase in population EVPI. Moreover, the effective life time of a new innovation is highly uncertain. Furthermore, the choice to use national or global estimates of the annual incidence of the disease is rather arbitrary, but even more influential for the population EVPI.

Because the central limit theorem held for all parameters, we could use normal distributions to represent uncertainty. This enabled us to introduce correlations between parameters, drawing from the multivariate normal distribution. Not taking correlations into account can result in an over- or underestimation of the EVPI (see appendix).

Interestingly, with value of information analysis the question arises again whether we care about the distribution of the outcome. Figure 3 illustrates that the value of information is highly uncertain: there is a 61% probability that it is zero, it is expected to be 83 euro, and the probability that it exceeds 200 euro per patient is 17%. Expected utility theory would again state that we should only care about the 83 euro. Nevertheless, the risk attitude of the researcher and grant provider could influence the choice between two projects with the same population EVPI. The more skewed the distribution of the value of information, the more likely a future study will show no benefit. However, skewness makes it also more likely to find that another strategy is actually substantially better: a less likely but very important finding.

Value of information analysis is an elegant technique to answer an essential question: should we perform more research? It has been receiving increasing attention in the literature, but only a few applications have been published(3). We illustrated the practicality and usefulness of value of information analysis to assess the expected value of more research and identify research priorities concerning the use of MRI for acute knee trauma. More research seems justified, although EVSI analysis should confirm this, and the focus should be on only 4 out of 108 parameters.

References

1. Claxton, K., The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. J Health Econ, 1999. 18(3): p. 341-64.
2. Claxton, K. and J. Posnett, An economic approach to clinical trial design and research priority-setting. Health Econ, 1996. 5(6): p. 513-24.
3. Claxton, K., et al., Bayesian value-of-information analysis. An application to a policy model of Alzheimer's disease. Int J Technol Assess Health Care, 2001. 17(1): p. 38-55.
4. Phillips, C.V., The economics of 'more research is needed'. Int J Epidemiol, 2001. 30(4): p. 771-6.
5. Oostenbrink, J.B., M.A. Koopmanschap, and F.F.H. Rutten, Handleiding voor kostenonderzoek, methoden en richtlijnprijzen voor economische evaluaties in de gezondheidszorg. 2000, Amstelveen: College voor zorgverzekeringen.
6. Hunink, M.G.M. and P.P. Glasziou, Decision making in health and medicine - integrating evidence and values. 2001, Cambrige: Cambridge University Press.

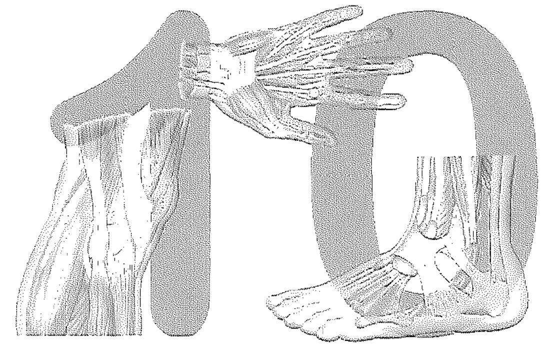
7. Gold, M.R., et al., Cost-effectiveness in health and medicine. 1996, Oxford: Oxford University Press.
8. Koopmanschap, M.A., et al., The friction cost method for measuring indirect costs of disease. J Health Econ, 1995. 14(2): p. 171-89.
9. Critchfield, G.C., K.E. Willard, and D.P. Connelly, Probabilistic sensitivity analysis methods for general decision models. Comput Biomed Res, 1986. 19(3): p. 254-65.
10. Doubilet, P., et al., Probabilistic sensitivity analysis using Monte Carlo simulation. A practical approach. Med Decis Making, 1985. 5(2): p. 157-77.
11. Briggs, A.H., et al., Probabilistic analysis of cost-effectiveness models: choosing between treatment strategies for gastroesophageal reflux disease. Med Decis Making, 2002. 22(4): p. 290-308.
12. Brennan, A., et al., A two level monte carlo approach to calculating expected value of perfect information:resolution of the uncertainty in methods.[abstract]. Med Decis Making, 2002. 22(6): p. 552.
13. Ades, A.E., G. Lu, and K. Claxton, Expected Value of Sample Information calculations in medical decision making. Med Decis Making, 2003. submitted.
14. Felli, J.C. and G.B. Hazen, Sensitivity analysis and the expected value of perfect information. Med Decis Making, 1998. 18(1): p. 95-109.
15. Raiffa, H., Decision Analysis: Introductory Lectures on Choices under Uncertainty. Reading, MA: Addison-Wesley, 1961.
16. Weinstein, M.C., et al., Clinical decision analysis. 1980, Philadelphia: W.B. Saunders Company.

Appendix

Imagine y is the outcome and x_1 and x_2 are two distributions representing resource use. If $y = x_1 + x_2$, then the general formula for the variance of y is:

$$var(y) = var(x_1) + var(x_2) + \rho * sd(x_1) * sd(x_2).$$

This formula shows that the effect of disregarding the correlation (ρ) depends on the sign of the correlation. Since in our model most correlations were positive, disregarding them led to smaller distributions and an underestimation of the total EVPI.



Summary and Discussion

Summary

In search for improvement of the early diagnostic accuracy in patients presenting with acute trauma of wrist, knee, or ankle, we studied the value of a short low-field MRI examination in addition to radiography in a randomized controlled trial (RCT). We were interested to know whether the addition of a short MRI examination in all patients with acute trauma of wrist, knee, and ankle would improve our ability to discriminate between patients requiring additional treatment after their initial hospital visit and patients that could be sent home without further follow-up. We analyzed the effects and the costs of the addition of the short MRI examination to the current diagnostic strategy.

Meta-analysis

Chapter 2 describes a meta-analysis of the published data on the diagnostic performance of magnetic resonance imaging of the menisci and cruciate ligaments. In this analysis we also assessed the effect of study design characteristics and magnetic field strength on diagnostic performance. Articles on the diagnostic performance of MRI of knee lesions published between 1991 and 2000 were included if at least 30 patients were studied, arthroscopy was the reference standard, the magnetic field strength of the MRI unit was reported, positivity criteria were defined, and the absolute numbers of true-positive, false-negative, true-negative, and false-positive results were available or derivable. Pooled-weighted and summary receiver operating characteristic (SROC) analyses were performed for tears of both menisci and both cruciate ligaments separately and for the four lesions combined, using random effects models. Differences across lesions were assessed, as was the effect of magnetic field strength and characteristics of study design on diagnostic performance.

Twenty-nine of 120 retrieved articles were included. Pooled sensitivity was higher for medial meniscal tears than for lateral meniscal tears. However, pooled specificity for the medial meniscus was lower than for the lateral meniscus. In SROC analyses performed per lesion, various study design characteristics were found to influence diagnostic performance. Higher magnetic field strength significantly improved discriminatory power only for anterior cruciate ligament tears. When all lesions were combined in one overall SROC analysis, magnetic field strength was a significant but modest predictor of diagnostic performance.

We concluded that the diagnostic performance of knee joint MRI was different across lesions and was influenced by various study design characteristics. Higher magnetic field strength modestly improved diagnostic performance, but a substantial effect could only be demonstrated for anterior cruciate ligament tears.

Development and evaluation of short MRI scanning protocols

In **chapter 3** the development of a short MRI scanning protocol to be used in the RCT assessing the costs and effects of a short MRI examination in patients with acute peripheral joint injury is described. We compared the short protocol with standard MRI protocols as currently in use.

Various parameters of sequences frequently used for musculoskeletal MRI were adapted in order to decrease scanning time while preserving an acceptable diagnostic quality. The four sequences that appeared most suitable (GE-T1w, SE-T1w, Turbo Dual Echo, and STIR) for the intended purpose were compared to standard MRI examination in 80 patients with peripheral joint trauma.

The short MRI protocol as a screening examination for the presence or absence of pathology had a sensitivity of 0.99 and a specificity of 0.90. Fracture and bone marrow edema could be demonstrated accurately with a sensitivity of 1.00 and 0.91 respectively and a specificity of 1.00 for both. The sensitivity and specificity for anterior cruciate ligament rupture was 0.78 and 0.90 respectively. For meniscus tears the specificity was high (medial 0.95, lateral 0.94), but the sensitivity was lower (medial 0.73, lateral 0.58).

We concluded that the short MRI protocol couldn't replace the standard protocol for the comprehensive examination of wrist, knee, and ankle for all indications, but that it may be suitable as a screening tool in the initial investigation of patients with peripheral joint trauma.

Prediction rules

In **chapter 4, 5, and 6** the value of a short MRI examination in all patients with acute trauma of wrist, knee or ankle was assessed in a randomized controlled trial with the purpose of identifying patients who need additional treatment versus those patients that do not need therapy and may be sent home without follow-up.

Ninety patients with wrist trauma, 208 patients with knee trauma, and 202 patients with ankle trauma were randomized between radiography and radiography followed by a short MRI examination on a low field dedicated extremity MRI system. Data were collected on sex, age, trauma mechanism, radiography result, MRI result, and treatment. For the wrist also tenderness of the anatomical snuffbox was recorded. Univariable and multivariable logistic regression analysis was used to create four models for each joint predicting treatment after the initial hospital visit. The MRI result was analyzed as either positive or negative compared to 'no MRI information available' as reference group. Model performance was assessed using Akaike's Information Criterion and ROC analysis.

Of the 90 patients with wrist injury three patients were excluded. Thirty-six fractures were found, of which seven were radiographically occult (four of the distal radius and three of the scaphoid). Only one patient had a significant soft tissue lesion consisting of a ruptured tendon. In univariable analysis age (likelihood ratio p-value 0.07), snuffbox tenderness (p-value 0.08), radiography result (p-value<0.01), and MRI result (p-value 0.06) significantly predicted treatment, using a likelihood ratio p-value of 0.1. In the multivariable analysis the contribution of snuffbox tenderness was not significant (using a p-value threshold of 0.05). Radiography result and MRI result both significantly contributed in predicting the need for treatment. However, only a positive MRI result (compared to no MRI made) had significant predictive value, whereas a negative result did not.

Of the 208 patients with knee injury 16 were excluded, mainly because on reassessment not all inclusion criteria were fulfilled. Treatment was required in 111 patients, consisting of arthroscopy or open surgery, immobilization and physiotherapy in 27, 25 and 59 cases respectively. Age 30 years or higher, indirect trauma mechanism, radiography result and MRI result were significant predictors of treatment in the univariable analysis (likelihood ratio p-value 0.01, <0.001, 0.02 and <0.001 respectively). In the multivariable analysis MRI contributed significantly to the prediction model (likelihood ratio test p-value 0.04), but the models with MRI did not discriminate better between patients in need of additional treatment and those not requiring treatment than the model without MRI. The odds ratios of the MRI result in addition to radiography were not significant, neither for a positive nor for a negative result.

Of the 202 patients with ankle injury, five were excluded; 109 of the 197 remaining patients received additional treatment, consisting of surgery, plaster immobilization, softcast, tape, or physiotherapy. In univariable analysis age (likelihood ratio test p-value 0.08), radiography result (p-value <0.001) and MRI result (p-value 0.002) had a likelihood ratio p-value less than 0.1 and were subsequently analyzed in multivariable regression analyses. In the multivariable analysis, a positive MRI result significantly improved the prediction of the need for treatment, but a negative result did not predict that treatment was not necessary. The area under the ROC curve of the model with both radiography and MRI result did not differ significantly from the model with only radiography, indicating no difference in discriminatory power.

We concluded that a short MRI examination on a low field MRI system in the initial workup of all patients with acute trauma of wrist, knee or ankle has additional value in predicting the need for treatment in addition to radiography. However, in wrist and ankle trauma, we could only demonstrate that a positive MRI result had a significant additional predictive value; in knee trauma, the additional value of a positive or a negative result were both not significant. Therefore, the addition of a short MRI examination to the current workup can not improve discrimination between patients requiring therapy and those that can be sent home without treatment. In patients with wrist injury early MRI may be valuable when performed in patients clinically suspected of a scaphoid fracture but with normal radiographs, however, the number of patients in this subgroup was too small for meaningful analysis. In patients with ankle injury MRI in the initial evaluation may be valuable in a setting where ruptured ankle ligaments are immediately operated, which is not the case in our hospital.

Analysis of costs and effects

In **chapter** 7 the costs and effects of the short dedicated extremity MRI examination implemented in the initial diagnostic evaluation of all patients with acute injury of wrist, knee or ankle were studied. All costs associated with the initial injury were analyzed from the societal perspective. Secondary outcome measures were number of additional diagnostic procedures, time to last diagnostic procedure, number of days absent from work, time to convalescence, and quality of life, measured using the EuroQol and SF-36 health survey questionnaires at four points in time until 6 months after the injury.

For wrist injury, the total costs were highest in the intervention group (the group that underwent the additional short MRI examination). For ankle injury the total costs in the intervention group and the reference group were almost equal. The total costs in the intervention group of knee-injured patients were considerably lower than in the reference group, which was mainly attributable to friction costs. However, for all joints the difference in costs between the intervention groups and the reference groups was not statistically significant. In the intervention group of knee-injured patients, the number of diagnostic procedures was significantly lower and time to completion of the diagnostic workup was significantly shorter than in the reference group. In patients with wrist or ankle injury, non of these measures of effect differed significantly. Although for all three joints time to convalescence was lower in the intervention group, the difference with the reference group was not significant. In wrist and ankle-injured patients the EuroQol tariff measured one week, 6 weeks, 3 months and 6 months after the initial injury was at no point in

time significantly different between the intervention group and the reference group. In knee-injured patients the EuroQol tariff was significantly higher in the intervention group at one week and at 6 weeks after the initial trauma. After 3 months this difference was not significant anymore. The eight domain scores of the SF-36 questionnaire were only in a few occasions significantly different between the intervention group and the reference group. For the wrist the vitality and role-emotional domains scored higher in the intervention group at 6 weeks. For the knee, the physical functioning score was higher in the intervention group after one week. For the ankle, the mental health score was lower for the intervention group after 6 weeks and after 6 months. We concluded that a short MRI examination in acute wrist or ankle injury neither saves costs nor does it increase effectiveness. In knee-injured patients a short MRI examination shortens the time to completion of the diagnostic workup, reduces the number of diagnostic procedures, improves quality of life in the first 6 weeks, and the data suggest that it may reduce costs associated with lost productivity.
In **chapter** 8 we analyzed if the cost-effectiveness of the application of a short MRI examination can be increased by selecting patients without a fracture on the initial radiograph. This strategy was modeled from the intervention and the reference group and therefore the significance of differences could not be tested statistically. All measures of effect in the modeled group were more favorable than in the reference group but similar compared to the intervention group. However, the overall costs were lowest if MRI was performed only in patients who showed no fracture on the radiograph and highest if radiography was never followed by a short MRI examination. The results suggest that selective use of a short MRI examination following radiography in patients with acute knee injury who have no fracture on the radiograph saves costs and potentially increases effectiveness.
A value of information analysis is described in **chapter 9** and shows that additional research into the costs of a short MRI examination in knee-injured patients without a fracture on the initial radiograph is expected to show a further decrease in overall costs of 83 Euro per patient. Assuming the number of future patients who could potentially benefit from more research to be 20,000 per year for the Netherlands, the population EVPI was estimated to be 1.6 million Euro per year. This outcome warrants further research since the expected benefit of more research exceeds the expected costs of the research itself. Analysis of partial value of information showed that additional research should not focus on all variables that influence the overall costs, but that a better assessment of the time off work after an acute knee injury would suffice to assess the potential decrease of costs.

Discussion

We could not demonstrate that a short MRI examination in all patients with acute knee injury had additional value above the current diagnostic strategy in predicting the need for treatment after the initial hospital visit. In wrist and ankle injured patients the short MRI examination increased the predictive value, but this increase could only be attributed to the contribution of a positive MRI result. A negative MRI result did not add significantly in the prediction that the patient would not need additional treatment. Therefore, a short MRI examination is not useful as additional initial examination to decide which patient with acute joint trauma can be sent home without need of follow-up and additional treatment. Although we expected that an abnormal MRI result would add to the prediction of subsequent therapy in knee injured patients, rather than in wrist and ankle injury, the opposite appeared true. Our initial expectation was based on the fact that in daily routine MRI plays an important role in the detection and management of lesions of the knee, more than in wrist and ankle injury. However, it appeared that only 40 % of the patients with a meniscus rupture diagnosed by the initial MRI underwent partial meniscectomy within 6 months. Some of the patients that did not undergo arthroscopy despite a positive MRI examination for meniscus rupture still had complaints at the end of the follow-up period of 6 months, but others were free of symptoms within a few months. It is likely that some of the patients with prolonged complaints indeed had a meniscus rupture and, on the other hand, that some of the patients with complaints of short duration had either a false positive MRI result, or a spontaneously healed meniscus rupture. This observation implies that the early MRI examination, besides expediting necessary treatment, increases the risk of unnecessary treatment.

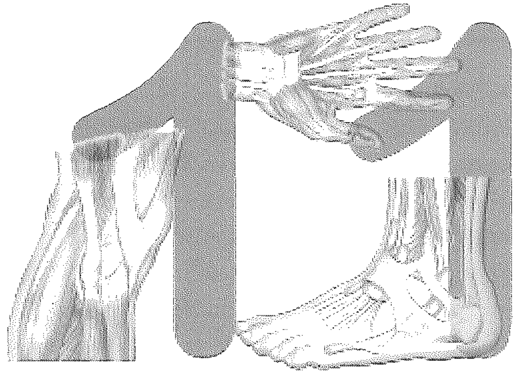
The fact that we could not prove that a short initial MRI examination had additional value in discriminating between patients that would and would not need additional treatment did not preclude a potential role of the application of MRI. The analysis of the costs and effects would likely have benefited from a strong additional value in discriminating patients, but did not depend on it, because the main potential cost reduction to society was from a reduction in time to final diagnosis, resulting in earlier treatment followed by earlier recovery and earlier resumption of work. However, the favorable effects of a short initial MRI examination in all patients with wrist and ankle injury were limited and the overall costs were not reduced. The effects in the subgroup of patients suspected from a scaphoid fracture were promising, but could not be evaluated due to the small number of patients with a scaphoid fracture.

Likewise, both costs and effects in patients with acute knee injury were promising, but a significant difference in costs could not be proven due to the wide variation in days off work and therefore in friction costs.

We demonstrated that the reduction in total costs is likely larger if only patients without a fracture on the radiograph are selected for MRI examination. One could argue that, from a societal perspective, it would be preferable to apply the short initial MRI examination only in patients with acute knee injury who are economially productive, since the potential cost-reduction is mainly caused by a reduction in lost productivity. This strategy would almost certainly result in an increase in cost-effectiveness and a further decrease in total costs to society. However, non-workers would likely feel discriminated against, since they would not benefit directly. Their acceptance of such a strategy would likely be low, even though their access to health care would not decrease and society as a whole would benefit by lower costs.

The mean total costs in the model for the value of information (VOI) analysis differ from these values in the RCT. The reason is that in the model for the VOI analysis we assumed the friction costs per day to be equal for all randomized subjects, whereas in the RCT we were bound to the data as obtained. In the RCT differences in friction costs between the groups were influenced by differences in age and sex distribution between the groups, since friction costs were calculated using estimated age and sex specific friction costs per day. In the VOI-model we assumed that differences in age and sex distribution between strategies could be attributed to chance only and therefore we calculated mean friction costs per day over all randomized subjects.

From a decision-analytic point of view the results of the study would lead to the conclusion that a short MRI examination should be performed in all patients with an acute knee injury, who show no fracture on the radiograph. The fact that the advantage of this strategy has not been shown with statistical significance does not preclude this conclusion, since it is the best available evidence. Because the advantage of the new strategy is mainly an expected cost reduction on the societal level and not on the hospital or the third party payers level, the decision to implement the strategy should be taken at the government level and should be accompanied by some form of reimbursement. Since decision-makers are generally risk aversive, the question arises if the uncertainty about the results should be decreased by additional research and to what extent the additional research is expected to affect the outcome. In a value of information analysis, we demonstrated that further research into the cost of a short MRI examination in patients with acute knee injury is justified and is expected to exceed the costs of the research. Additional research may be limited to an RCT, assessing the number of days off work for the different strategies.



Samenvatting en Discussie

Samenvatting

Bij acuut letsel van pols, knie of enkel bestaat het eerste medische onderzoek doorgaans uit anamnese en lichamelijk onderzoek, vaak aangevuld met een röntgenfoto. Frakturen kunnen met deze strategie meestal goed worden aangetoond, maar de aanwezigheid en de uitgebreidheid van weke delen letsel, zoals een gescheurd ligament of een meniscusscheur, zijn over het algemeen niet makkelijk aan te tonen. Dit wordt mede veroorzaakt doordat in de acute fase het gewricht pijnlijk is en daardoor vaak minder goed te onderzoeken. MRI zou hier een potentiële rol kunnen vervullen omdat met MRI veel weke delen letsels goed te visualiseren zijn. Een belangrijke reden waarom MRI tot nu toe tochvrijwel niet toegepast wordt bij het eerste onderzoek van acuut gewrichtsletsel, is gelegen in de hoge kosten en de lange onderzoeksduur van MRI. Daarnaast is nog niet geheel duidelijk in hoeverre de informatie die verkregen zou worden, als MRI in het acute stadium zou worden toegepast, invloed heeft op behandeling en uiteindelijke uitkomst voor de patiënt.

De ontwikkeling van steeds betere en relatief goedkope laagveld dedicated MRI scanners in het afgelopen decennium heeft er toe geleid dat MRI nu mogelijk wel kosteneffectief toegepast kan worden

In dit proefschrift wordt onderzocht wat de invloed is van de toepassing van een kort MRI onderzoek met een laag-veld MRI scanner (veldsterkte 0,2 Tesla) bij het eerste onderzoek van patiënten met een acuut letsel van pols, knie of enkel.

In **hoofdstuk 2** wordt in een meta-analyse van 29 artikelen, die de diagnostische performance van MRI van knieletsels beschrijven, de samengestelde gewogen sensitiviteit en specificiteit van MRI voor mediale en laterale meniscusrupturen en voorste en achterste kruisbandrupturen onderzocht. In een 'summary receiver operating characteristic' (SROC) analyse wordt de invloed van diverse studie karakteristieken op de diagnostische performance geanalyseerd. In de analyse wordt gebruik gemaakt van 'random effects modellen'. De samengestelde sensitiviteit blijkt voor de mediale meniscusscheuren hoger dan voor laterale meniscusscheuren, terwijl de samengestelde specificiteit voor laterale meniscusscheuren hoger is dan voor mediale meniscusscheuren. Uit de SROC analyse blijkt dat hogere veldsterkte de diagnostische performance alleen voor voorste kruisbandrupturen significant verbetert. Wanneer alle laesies bij elkaar genomen worden in een 'overall SROC analyse' blijkt dat hogere veldsterkte een beperkte, maar significante verbetering van diagnostische performance laat zien.

In **hoofdstuk** 3 wordt de ontwikkeling en evaluatie van een korte MRI protocollen voor onderzoek van pols, knie en enkel beschreven. Het doel van deze korte MRI protocollen is het verkorten van de onderzoekstijd en daarmee tevens het beperken van de kosten. Diverse parameters van sequenties die geschikt zijn voor muskuloskeletale MRI gebruikt werden hiervoor aangepast. Dit leidde onvermijdelijk tot enige achteruitgang in kwaliteit van de afbeeldingen. De parameters werden dusdanig aangepast dat subjectief nog een acceptabele afbeeldingkwaliteit werd verkregen. Een protocol van vier verkorte sequenties (T1-gewogen spin echo, T1-gewogen gradiënt echo, turbo dual echo en STIR) werd vergeleken met een standaard protocol dat gebruikelijk is voor muskuloskeletale MRI, om een indruk te krijgen in hoeverre het aanpassen van het protocol van invloed was op de diagnostische performance. Vergeleken met het standaard protocol had het korte protocol een sensitiviteit van 0,99 en een specificiteit van 0,91 voor het aantonen van de aan- of afwezigheid van pathologie. Voor frakturen en beenmergoedeem werd een sensitiviteit van 1,00 respectievelijk 0,91 en een specificiteit van 1,00 verkregen. De sensitiviteit en specificiteit voor voorste kruisbandruptuur was 0,78 respectievelijk 0.90. Voor meniscusscheuren werd een hoge specificiteit verkregen (mediaal 0,95, lateraal 0,94), maar een lagere sensitiviteit (mediaal 0,73, lateraal 0,58). We concludeerden dat de korte MRI protocollen de standaard protocollen niet konden vervangen voor iedere indicatie. De kort protocollen hebben mogelijk wel een meerwaarde, indien gebruikt als eerste onderzoeksmodaliteit bij de evaluatie van acuut perifeer gewrichtsletsel.

Predictieregels

In **hoofdstuk 4, 5 en 6** wordt de waarde van een kort MRI onderzoek bij alle patiënten met een acuut letsel van pols, knie of enkel onderzocht in een gerandomiseerd gecontroleerde studie, met als doel een onderscheid te kunnen maken tussen patiënten die aanvullende behandeling nodig hebben na hun eerste bezoek aan het ziekenhuis versus patiënten die geen aanvullende behandeling nodig hebben en dus ook geen polikliniekafspraak ter controle hoeven te krijgen. Negentig patiënten met polsletsel, 208 patiënten met knieletsel en 202 patiënten met enkelletsel werden gerandomiseerd tussen röntgenonderzoek en röntgenonderzoek gevolgd door een kort MRI onderzoek met een laagveld 'dedicated extremity' MRI systeem. Gegevens betreffende geslacht, leeftijd, trauma mechanisme, röntgenuitslag, MRI uitslag en behandeling werden verzameld. Voor de pols werd ook de aanwezigheid van drukpijn in het tabatière anatomique genoteerd.

Met behulp van univariabele en multivariabele logistische regressie analyse werden voor ieder gewricht vier modellen gecreëerd voor de predictie van behandeling na het eerste ziekenhuisbezoek. De performance van de modellen werd beoordeeld met Akaike's Information Criterium en ROC analyse.

Van de 90 patiënten met een polsletsel werden er drie geëxcludeerd. Zesendertig patiënten hadden een fraktuur, waarvan zeven alleen met MRI zichtbaar. Slechts één patiënt had een significant weke delen letsel, te weten een peesruptuur. In univariabele logistische regressie analyse werden leeftijd, drukpijn in het tabatière anatomique, röntgen uitslag en MRI uitslag geselecteerd voor multivariabele analyse. In de multivariabele logistische regressie analyse bleek de bijdrage van drukpijn in het tabatière anatomique niet significant bijdragend. Röntgen uitslag en MRI uitslag droegen beiden significant bij in de predictie van behandeling. Echter, als het ontbreken van MRI informatie als referentie werd beschouwd (de huidige situatie), was alleen de bijdrage van een positieve MRI uitslag significant bijdragend, terwijl een negatieve MRI uitslag geen aanvullende predictieve waarde had.

Van de 208 patiënten met knieletsel werden 16 geëxcludeerd, voornamelijk omdat na inclusie bleek dat niet aan alle inclusievoorwaarden was voldaan. In univariabele logistische regressie analyse werden leeftijd, trauma mechanisme, röntgen uitslag en MRI uitslag geselecteerd voor nadere multivariabele analyse. Al deze variabelen droegen ook significant bij in multivariabele regressie analyse aan de predictie van behandeling. Echter, als een negatieve of een positieve MRI uitslag werd geanalyseerd met als referentie het ontbreken van MRI informatie, was de aanvullende waarde van ieder van deze uitslagen niet significant.

Van de 202 patiënten met enkelletsel werden vijf geëxcludeerd. In univariabele logistische regressie analyse werden leeftijd, röntgen uitslag en MRI uitslag geselecteerd voor nadere multivariabele analyse. Alle variabelen droegen ook significant bij in de multivariabele analyse, maar vergeleken met het ontbreken van MRI informatie droeg alleen een positieve MRI uitslag significant bij aan de predictie van behandeling, terwijl een negatieve uitslag niet bijdroeg.

Voor alle gewrichten verschilde het oppervlak onder de ROC curve van het model zonder MRI niet significant van de oppervlakten onder de ROC curve van de modellen met MRI, hetgeen impliceert dat er geen significant verschil in discriminerend vermogen tussen de modellen aangetoond kon worden.

We concludeerden dat een kort MRI onderzoek met een laagveld 'dedicated extremity' MRI systeem bij het eerste onderzoek van patiënten met acuut letsel van pols, knie of enkel aanvullende waarde heeft voor de predictie van behandeling na het eerste ziekenhuisbezoek. Echter, bij pols- en enkelletsel kon, indien vergeleken werd met het ontbreken van MRI informatie, alleen voor een positief MRI resultaat een significant aanvullende waarde worden aangetoond. Bij knieletsel was de bijdrage van zowel een positieve als een negatieve MRI uitslag, met als referentie het ontbreken van MRI informatie, niet significant bijdragend. De toevoeging van een kort MRI onderzoek aan de huidige diagnostische strategie verbetert dus niet het vermogen onderscheid te maken tussen patiënten die aanvullende behandeling nodig hebben en patiënten die naar huis gestuurd kunnen worden, zonder noodzaak van aanvullende controle.
Bij patiënten die verdacht worden van een scaphoid fractuur terwijl de röntgenfoto geen fraktuur laat zien is het korte MRI onderzoek mogelijk waardevol, echter de aantallen waren te klein voor analyse van deze subgroep. Bij patiënten met enkelletsel kan een kort MRI onderzoek waardevol zijn in een setting waarbij de behandeling van gescheurde enkelligamenten in eerste instantie operatief is. Dit is in Nederland echter niet gebruikelijk, zodat dit niet geanalyseerd kon worden.

Analyse van kosten en effekten.
Hoofdstuk 7 beschrijft een analyse van de kosten en effectiviteit van een kort MRI onderzoek als initieel diagnostische modaliteit bij akuut letsel van pols, knie en enkel. Alle kosten, geassocieerd met het letsel werden geanalyseerd vanuit maatschappelijk perspektief. Secundaire uitkomstmaten waren aantal aanvullend diagnostische procedures, tijd tot laatste diagnostische procedure, aantal dagen werkverzuim, tijd tot herstel en kwaliteit van leven. Kwaliteit van leven werd gemeten met de EuroQol en SF-36 op vier momenten gedurende een follow-up periode van zes maanden.
Bij polsletsel waren de totale kosten het hoogst in de groep die een kort MRI onderzoek had ondergaan (gemiddelde kosten 1106 Euro versus 905 Euro in de groep die geen MRI had ondergaan). Bij enkelletsel waren de totale kosten vrijwel gelijk in de groep die een kort MRI onderzoek had ondergaan en de groep die alleen röntgenonderzoek had ondergaan (gemiddelde kosten 1675 Euro versus 1660 Euro). De totale kosten van patiënten met een knieletsel waren aanzienlijk lager in de groep die een kort MRI onderzoek had ondergaan dan in de groep die geen kort MRI onderzoek had ondergaan (gemiddelde kosten 1820 Euro versus 2231 Euro).

Dit verschil in totale kosten werd vrijwel geheel veroorzaakt door een verschil in frictiekosten (maatschappelijke kosten veroorzaakt door produktiviteitverlies). Voor alle drie de gewrichten waren de gevonden verschillen in kosten echter niet significant. Bij patiënten met een knieletsel was in de groep die een kort MRI onderzoek had ondergaan de kwaliteit van leven, gemeten met de EuroQol, gedurende de eerste 6 weken significant beter, de tijdsduur tot het laatste diagnostische onderzoek significant korter en het aantal diagnostische onderzoeken significant minder vergeleken met de groep die geen kort MRI onderzoek had ondergaan. De duur van het werkverzuim en van de tijd tot herstel verschilden niet significant. Bij patiënten met pols- of knieletsel werd in geen van deze effectiviteitsmaten een significant verschil gevonden tussen de groep die wel en de groep die geen kort MRI onderzoek had ondergaan.
Concluderend vonden we bij patiënten met een pols- of enkelletsel geen significante reductie in kosten noch een toename in effectiviteit indien een kort MRI onderzoek werd verricht. Bij patiënten met een knieletsel leidde het korte MRI onderzoek tot een verkorting van tijd tot het laatste diagnostische onderzoek, een vermindering van het aantal diagnostische procedures en een verbetering van de kwaliteit van leven in de eerste zes weken. Verder geven de data een indicatie dat bij patiënten met een knieletsel het korte MRI onderzoek mogelijk leidt tot een reductie van kosten veroorzaakt door een vermindering van productiviteitverlies.
In **hoofdstuk** 8 wordt onderzocht of de kosteneffectiviteit van een kort MRI onderzoek bij patiënten met een acuut knieletsel kan worden verhoogd door selectieve toepassing bij patiënten die geen zichtbare fraktuur op de röntgenfoto hebben. Deze strategie werd geanalyseerd met een modelmatige benadering en derhalve konden verschillen niet statistisch getest worden. De patiëntengroep voor deze strategie werd gemodelleerd uit de twee gerandomiseerde groepen. Alle effectiviteitmaten, zoals ook gebruikt in hoofdstuk 7, waren gunstiger in de gemodelleerde groep vergeleken met de groep die alleen röntgenonderzoek had ondergaan en vrijwel gelijk aan de groep waarbij alle patiënten MRI onderzoek hadden ondergaan. De totale kosten waren echter het laagst indien het korte MRI onderzoek alleen werd verricht bij patiënten die op de röntgenfoto geen zichtbare fraktuur hadden. De totale kosten waren het hoogst in de groep patiënten die geen kort MRI onderzoek had ondergaan. De resultaten suggereren dat selectieve toepassing van een kort MRI onderzoek voor patiënten met een knieletsel, waarbij op de röntgenfoto geen fraktuur zichtbaar is, een verdere toename van de kostenbesparing geeft vergeleken met MRI bij alle patiënten met knie letsel.

Een 'value of information' analyse wordt beschreven in **hoofdstuk** 9 en toont aan dat aanvullend onderzoek naar de kosten van een kort MRI onderzoek bij patiënten met knieletsel, waarbij op de röntgenfoto geen fraktuur zichtbaar is, een verwachtte verdere vermindering van de totale kosten met 83 Euro zal laten zien. Indien het aantal patiënten in Nederland dat in de toekomst baat zal hebben van aanvullend onderzoek geschat wordt op 20.000 per jaar is de geschatte verwachtte waarde van perfecte informatie (Expected Value of Perfect Information - EVPI) 1,6 miljoen Euro per jaar. Deze uitkomst rechtvaardigt verder onderzoek aangezien de verwachtte opbrengst van meer onderzoek groter is dan de verwachtte kosten van verder onderzoek. Een 'partial value of information analyse' laat zien dat het niet noodzakelijk is alle variabelen die de totale kosten beïnvloeden nauwkeuriger te bepalen, maar dat een nauwkeurigere bepaling van het werkverzuim, veroorzaakt door het knieletsel, voldoende zal zijn om de potentiële kostenvermindering te analyseren.

Discussie

We hebben niet kunnen aantonen dat een kort MRI onderzoek bij alle patiënten met acuut knieletsel aanvullende waarde heeft boven de huidige strategie voor de predictie van behandeling na het eerste ziekenhuisbezoek. Bij patiënten met pols- en enkelletsel verhoogde het korte MRI onderzoek de predictieve waarde, maar deze toename in voorspellend vermogen werd alleen veroorzaakt door de bijdrage van een positieve MRI uitslag. Een negatieve MRI uitslag droeg niet significant bij aan de voorspelling dat de patiënt geen aanvullende behandeling nodig heeft. Het korte MRI onderzoek is dus niet nuttig om te differentiëren welke patiënt naar huis gestuurd kan worden zonder noodzaak tot verder onderzoek of behandeling. Onze verwachting was dat de voorspellende waarde van een afwijkende MRI uitslag, met betrekking tot behandeling na het eerste ziekenhuisbezoek, groter zou zijn voor knieletsel dan voor pols- of enkelletsel. Het tegenovergestelde bleek echter waar. Onze initiële verwachting was gebaseerd op het feit dat MRI in de dagelijkse praktijk bij knieletsel een belangrijke rol speelt, meer dan bij pols- of enkelletsel. Het bleek echter dat slechts 40% van de patiënten, bij wie het korte MRI onderzoek een meniscusruptuur liet zien, een partiele meniscectomie onderging in de erop volgende zes maanden. Sommige van deze patiënten hadden na zes maanden nog steeds klachten, maar anderen waren in enkele maanden klachtenvrij. Het is aannemelijk dat sommige van de patiënten met langdurige klachten inderdaad een meniscusruptuur hadden en dat sommige van de patiënten met kortdurende klachten een vals-positieve MRI uitslag hadden, dan wel een spontaan herstel van de meniscusruptuur.

Deze observatie impliceert dat het korte MRI onderzoek, naast het potentieel versnellen van noodzakelijke behandeling, het risico vergroot van onnodige behandeling.
Het feit dat we geen aanvullende waarde van het korte MRI onderzoek konden aantonen voor het differentiëren tussen patiënten die wel en die geen aanvullende behandeling nodig hebben sloot een potentiële rol voor de toepassing van MRI niet uit. De uitkomst van de kosten en effectiviteit analyse zou waarschijnlijk wel in positieve zin zijn beïnvloed als het korte MRI onderzoek een duidelijke aanvullende predictieve waarde zou hebben, maar noodzakelijk was dit niet, omdat de belangrijkste potentiële kostenreductie verwacht werd van een reductie in tijd tot diagnose, resulterend in een eerdere behandeling, gevolgd door een sneller herstel en een snellere hervatting van werk. De gunstige effekten van een kort MRI onderzoek bij alle patiënten met pols- en enkelletsel bleken echter beperkt en de totale kosten werden niet gereduceerd. De effekten in de subgroep van patiënten die verdacht werden van een scaphoidfraktuur waren veelbelovend, maar konden niet statistisch geanalyseerd worden omdat hun aantal te klein was. Bij patiënten met een acuut knieletsel werd een gunstige invloed gezien op zowel de kosten als de effecten, maar de kostenvermindering kon niet statistisch significant worden aangetoond, doordat er een grote variatie bestond in het werkverzuim en dientengevolge in frictiekosten.
We toonden aan dat de reductie in totale kosten waarschijnlijk nog sterker uitvalt als het korte MRI onderzoek alleen toegepast wordt bij patiënten die op de röntgenfoto geen zichtbare fraktuur hebben. Men zou kunnen stellen dat het vanuit maatschappelijk perspektief wenselijk is een kort MRI onderzoek alleen toe te passen bij patiënten die economisch productief zijn, aangezien de potentiële kostenreductie voornamelijk veroorzaakt wordt door een reductie van productiviteitsverlies. Een dergelijke strategie zou vrijwel zeker een verdere stijging van de kosteneffectiviteit tot gevolg hebben door een verdere daling van kosten voor de maatschappij. Het is echter zeer aannemelijk dat niet-werkenden zich door deze strategie gediscrimineerd zullen voelen, aangezien zij hooguit indirekt van de strategie kunnen profiteren. Hun acceptatie van deze strategie zal waarschijnlijk gering zijn, ondanks dat de toegankelijkheid van de gezondheidszorg voor hun niet zou afnemen en de totale kosten voor de maatschappij zouden verminderen.

De gemiddelde kosten in het model voor de value of information (VOI) analyse verschilde van de gemiddelde kosten in de gerandomiseerde studie. De oorzaak is dat we bij het model voor de VOI analyse aannamen dat de frictiekosten per dag hetzelfde waren voor alle gerandomiseerde deelnemers, terwijl we in de gerandomiseerde studie gebonden waren aan de gegevens zoals zij zich presenteerden.

In de gerandomiseerde studie werden verschillen in frictiekosten tussen de groepen beïnvloed door verschillen in de leeftijd-en geslachtdistributie tussen de groepen, aangezien frictiekosten werden berekend met behulp van leeftijds-en geslachtsafhankelijke frictiekosten per dag. In het VOI model namen we aan dat verschillen in leeftijd-en geslachtdistributie tussen de strategieën uitsluitend door toeval werden veroorzaakt en gebruikten we gemiddelde frictiekosten per dag voor alle gerandomiseerde patiënten.

Vanuit een besliskundig standpunt gezien leiden de resultaten van de studie tot de conclusie dat een kort MRI onderzoek verricht moet worden bij patiënten met een acuut knieletsel, waarbij op de röntgenfoto geen fraktuur zichtbaar is. Het feit dat het voordeel van deze strategie niet statistisch significant bewezen is, is vanuit besliskundig gezichtspunt niet van belang; de beslissing wordt gemaakt op basis van de best beschikbare gegevens. Aangezien het voordeel van de nieuwe strategie voornamelijk een verwachtte kostenreductie vanuit maatschappelijk perspektief is, en niet vanuit het perspektief van ziekenhuis of ziektekostenverzekeraar, zou de beslissing tot implementatie van de strategie genomen moeten worden op overheidsniveau en moeten samengaan met een vorm van vergoeding. Aangezien beleidsmakers in het algemeen risicomijdend zijn, komt de vraag op of de onzekerheid betreffende de resultaten verminderd zou moeten worden door aanvullend onderzoek en wat de verwachting is van de grootte van de verdere kostenreductie. In een VOI analyse toonden we aan dat verder onderzoek naar de kosten van een kort MRI onderzoek bij patiënten met een acuut knieletsel gerechtvaardigd is en dat de verwachtte verdere kostenreductie opweegt tegen de kosten van het aanvullende onderzoek. Aanvullend onderzoek kan beperkt blijven tot een gerandomiseerde studie, waarbij alleen het ziekteverzuim per strategie nader wordt onderzocht.

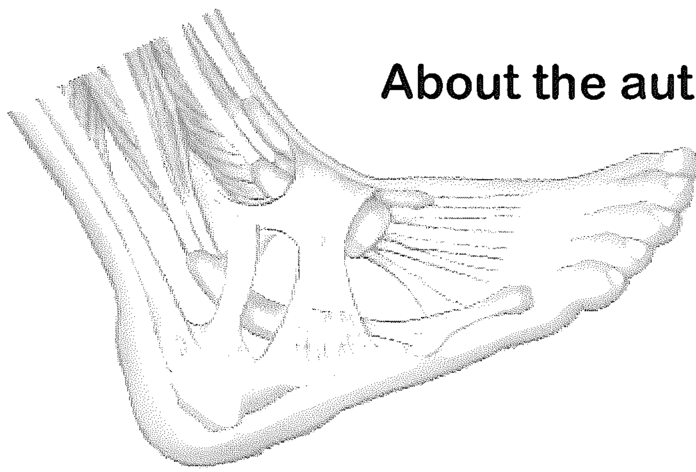
Dankwoord

Dankwoord

Velen hebben een aandeel gehad in het tot stand komen van dit proefschrift. Een aantal wil ik met name bedanken.
In de eerste plaats wil ik professor M.G.M. Hunink, mijn eerste promotor bedanken. Myriam, jouw grote gedrevenheid en kennis, gecombineerd met je zeer menselijke kant waren voor mij de beste mix als begeleiding van mijn onderzoek. Met de oprichting van de ART groep (Assessment of Radiological Technology) heb je gezorgd voor een omgeving waarin wetenschappelijk onderzoek goed kan gedijen. Vooral tijdens het laatste deel van mijn onderzoek heb ik daar dan ook optimaal gebruik van kunnen maken. De Art groep laat zien dat kruisbestuiving ook vele decennia na Mendel interessante wetenschappelijke resultaten kan opleveren.
Professor Krestin, mijn tweede promotor en initiator van het onderzoek beschreven in dit proefschrift, bedank ik voor de mogelijkheid die hij mij gegeven heeft dit promotieonderzoek te verrichten en voor zijn niet aflatende vertrouwen in een goede afloop.
Doctor A.Z. Ginai, mijn copromotor, bedank ik voor haar grote betrokkenheid bij en inzet voor het onderzoek. Ondanks haar fulltime job had zij altijd tijd, of maakte ze tijd, als ik langs kwam voor overleg of voor het beoordelen van MRI's - het zijn er vele honderden geweest.
Mijn paranimfen hebben beiden een belangrijk aandeel gehad in dit proefschrift. Edwin Oei, je onderzoeksopdracht in het kader van je Master of Science-opleiding is enigszins uit de hand gelopen. Jaren heb je intensief meegewerkt aan dit onderzoek; jouw bijdrage is onmisbaar geweest. Het doet me dan ook deugd dat je binnenkort de opleiding tot radioloog aan het Erasmus MC gaat beginnen. De diverse momenten van lichte wanhoop die ons tijdens het onderzoek af en toe bekropen hebben je er dus toch niet vanaf kunnen brengen.
Bas Groot Koerkamp, mijn tweede paranimf; ik denk met veel plezier aan onze gezamenlijke tijd op de ART kamer. Met onze vele discussies, zowel op wetenschappelijk als op ieder denkbaar ander vlak, had ik zo een paar hoofdstukken kunnen vullen met stellingen. Jou bijdrage is een verrijking van mijn proefschrift.
Professor J.A.N. Verhaar en professor A.B. van Vugt dank ik voor hun suggesties bij de opzet van het onderzoek en voor hun kritische commentaren op de manuscripten.
Andries Zwamborn, als Pater Familias van de afdeling Experimentele Radiologie ben jij bij de zwangerschap en geboorte van menig proefschrift betrokken geweest, zo ook bij dit proefschrift.

Dank voor je grote inzet, met name voor de lay-out en voor de prachtige uitwerking van de omslag, waarvoor je je met hart en ziel en enkele ledematen hebt ingezet. Teun Rijsdijk, dank voor je fotografie voor diverse van mijn artikelen. Ik zal altijd eerlijk de waarheid vertellen en zeggen dat de iets mindere kwaliteit van de afbeeldingen inherent is aan de studie opzet en niks van doen heeft met de kwaliteit van de fotograaf.
Wibeke van Leeuwen en Caroline van Bavel wil ik bedanken voor het vele werk dat zij hebben verzet met het invoeren van de vele duizenden gegevens in de database, het verzenden van enquêtes, de honderden telefoontjes naar patiënten die de enquête weer eens niet terugstuurden, enz. enz. Caroline heeft zich daarnaast ook de bediening van de Artoscan eigen gemaakt en heeft deel uitgemaakt van het team van laboranten dat gedurende bijna twee jaar naast hun gewone werk zich heeft ingezet voor het scannen van patiënten in weekenden en avonduren. Dit team bestond verder uit Baltus Alblas, Lourus van Haveren, Jolanda van het Kaar, Wim Lagendijk, Ton Louve en Lyanne Vollaard. Hen allen wil ik bedanken voor hun grote inzet. Verder wil ik alle andere laboranten bedanken die hebben meegeholpen met de inclusie van patiënten op de afdeling Spoed Eisende Hulp.
Professor Th. Stijnen en Karen Visser wil ik bedanken voor hun adviezen op methodologisch gebied.
Esaote dank ik voor hun support en voor de volledige wetenschappelijke vrijheid die mij daarbij gelaten is.
Tot slot wil ik Anja en mijn jongens, Tom, Menno en Michiel noemen; onbedoeld hebben jullie gezorgd voor het tot stand komen van dit proefschrift, maar tegelijkertijd zorgden jullie ervoor dat ik de relativiteit van al mijn inspanningen ben blijven zien, want jullie zijn pas echt belangrijk voor mij.

Jeroen Nikken.



About the author

Jeroen J. Nikken was born on October 12, 1964 in The Hague, The Netherlands. He attended the Gertrudislyceum in Roosendaal and graduated in 1983. He started his medical study at the University of Antwerp (RUCA). In 1984 he continued his medical study at the Erasmus University Rotterdam. After he finished his medical study, he worked on a primary healthcare project in Ojobi, Ghana. In 1991 he obtained his medical degree. Subsequently he accomplished his military service as an orthopaedic resident in the Central Military Hospital in Utrecht until 1993, followed by an orthopaedic residency in the Academic Hospital Utrecht. In 1994 he started working as a general surgery resident in the Refaja hospital in Dordrecht followed by the Havenziekenhuis in Rotterdam. In 1998 he started his research at the department of Experimental Radiology of the Erasmus University Rotterdam. In 2000 he started his training as a radiologist at the Erasmus MC (head of the department: Prof. dr. G.P. Krestin) in Rotterdam. He wrote the last part of his thesis at the department of Epidemiology and Biostatistics at the Erasmus MC, Rotterdam.