

# Abdominal MRI in women's health: advanced imaging

Maarten GJ Thomeer



# **Abdominal MRI in women's health: advanced imaging**

Maarten GJ Thomeer

**Abdominal MRI in women's health: advanced imaging**

Copyright 2018 © Maarten Guillaume Josephus Thomeer  
ISBN/EAN: 978-94-028-1231-2

Design and layout: Legatron Electronic Publishing, Rotterdam  
Printing: Ipskamp Printing, Enschede ([www.ipskampprinting.nl](http://www.ipskampprinting.nl))

One study in this thesis was financially supported by Nuts-Ohra.

Printing of this thesis was kindly supported by Bayer Healthcare.

No part of this thesis may be reproduced, stored in a retrieval system or transmitted in any form or by any means without permission from the author or, when appropriate, from the publishers of the publications.



# Abdominal MRI in women's health: advanced imaging

Gezondheid bij vrouwen:  
toegepaste MRI beeldvorming van de buik

Proefschrift

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

Prof.dr. R.C.M.E. Engels

en volgens besluit van het College voor Promoties.  
De openbare verdediging zal plaatsvinden op  
21 november 2018 om 15.30 uur

Maarten Guillaume Josephus Thomeer

geboren te Seria, Brunei

## **Promotiecommissie**

Promotor: Prof. Dr. M.G.M. Hunink

Copromotor: Dr. H.C. van Doorn

### **Leescommissie:**

Prof Dr. J. Stoker

Prof Dr. R. Beets-Tan

Prof Dr. J.N.M. IJzermans

# Contents

<b>Chapter 1</b>	General introduction	<b>1</b>
<b>PART I</b>	<b>GYNAECOLOGICAL IMAGING</b>	
<b>Chapter 2</b>	Clinical examination versus Magnetic Resonance Imaging for staging of cervical carcinoma: systematic review and meta-analysis <i>Maarten G Thomeer, Cees Gerestein, Sandra Spronk, Helena Van Doorn, Myriam Hunink</i> <i>Eur Radiol. 2013 Jul;23(7):2005–18</i>	<b>11</b>
<b>Chapter 3</b>	Evaluation of T2-W MR imaging and diffusion-weighted imaging for the early post-treatment local response assessment of patients treated conservatively for cervical cancer: a multicentre study <i>Maarten G Thomeer, Vincent Vandecaveye, Loes Braun, Frenchey Mayer, Martine Franckena-Schouten, Peter de Boer, Jaap Stoker, Erik Van Limbergen, Marrije Buist, Ignace Vergote, Myriam Hunink, Helena van Doorn</i> <i>Eur Radiol. 2018 Jun 25.</i>	<b>37</b>
<b>Chapter 4</b>	Can MR imaging at 3.0 Tesla reliably detect patients with endometriosis? Initial results <i>Maarten G Thomeer, Anneke B Steensma, Evert J van Santbrink, Francois E Willemsen, Piotr A Wielopolski, Myriam G Hunink, Sandra Spronk, Joop S Laven, Gabriel P Krestin</i> <i>J Obstet Gynaecol Res. 2014 Apr;40(4):1051-8</i>	<b>55</b>
<b>PART II</b>	<b>HEPATOCELLULAR ADENOMA</b>	
<b>Chapter 5</b>	Hepatocellular adenomas: correlation of MR imaging findings with pathologic subtype classification <i>Sanne van Aalten, Maarten Thomeer, Turkan Terkivatan, Roy Dwarkasing, Joanna Verheij, Rob de Man, Jan IJzermans</i> <i>Radiology. 2011 Oct;261(1):172-81.</i>	<b>71</b>
<b>Chapter 6</b>	MRI features of inflammatory hepatocellular adenomas on hepatocyte phase imaging with liver-specific contrast agents <i>Maarten G Thomeer, Francois E Willemsen, Katharina K Biermann, H Addouli, Rob A de Man, Jan J IJzermans, Roy S Dwarkasing</i> <i>J Magn Reson Imaging. 2014 May;39(5):1259-64</i>	<b>89</b>

<b>Chapter 7</b>	Genotype-phenotype correlations in hepatocellular adenoma: an update of MRI finding <i>Maarten G Thomeer, Mirelle E Bröker, Quido de Lussanet, Katharina Biermann, Roy Dwarkasing, Rob de Man, Jan N IJzermans, Marianne de Vries Diagn Interv Radiol. 2014 May-Jun;20(3):193-9.</i>	<b>103</b>
<b>Chapter 8</b>	Letter to the editor: Quantitative analysis of hepatocellular adenoma and focal nodular hyperplasia in the hepatobiliary phase: external validation of LLCER method using gadobenate dimeglumine as contrast agent <i>Maarten G Thomeer, Bibiche Gest, Hermen van Beek, Marianne De Vries, Roy Dwarkasing, Julia Klompenhouwer, Robert A De Man, Jan N IJzermans, Loes Braun J Magn Reson Imaging. 2018 Mar;47(3):860-861</i>	<b>117</b>
<b>Chapter 9</b>	Intra-patient comparison of the hepatobiliary phase of Gd-BOPTA and Gd-EOB-DTPA in the differentiation of HCA from FNH <i>Inge JSML Vanhooymissen, Maarten G Thomeer, Loes Braun, Bibiche Gest, Sebastiaan van Koeverden, Francois Willemsen, Myriam Hunink, Robert A De Man, Jan N IJzermans, Roy S Dwarkasing Accepted for J Magn Reson Imaging</i>	<b>121</b>
<b>Chapter 10</b>	Hepatocellular adenoma: when and how to treat? Update of current evidence <i>Maarten G Thomeer, Mirelle Broker, Joanne Verheij, Michael Doukas, Turkan Terkivatan, Diederick Bijdevaate, Robert A De Man, Adriaan Moelker, Jan N IJzermans Therap Adv Gastroenterol. 2016 Nov;9(6):898-912</i>	<b>139</b>
<b>Chapter 11</b>	Discussion Samenvatting	<b>163</b> 165 171
<b>Chapter 12</b>	List of publications PhD Portfolio Curriculum vitae Maarten Guillaume Josephus Thomeer Dankwoord	<b>175</b> 177 183 187 189

# Chapter 1

## General introduction

The World Health Organisation positions women's health within a wider body of knowledge that places importance on gender as a social determinant of health (<http://www.who.int/gender-equity-rights/en/>). In addition to social inequality, women are susceptible to a broad range of gender-specific diseases.

In the early twentieth century, death from uterine (uterine body and cervix) cancers was the leading cause of cancer death in women ([https://en.wikipedia.org/wiki/Women's\\_health](https://en.wikipedia.org/wiki/Women's_health)), but by the 1950's uterine cancer deaths had diminished significantly, mainly due to screening [1], and today cervical carcinoma is mainly a disease of developing countries. Nevertheless, morbidity and mortality from cervical cancer is still significant, even in industrialized countries, and in 2013 alone over 4000 deaths were attributable to cervical cancer in United States of America. Fortunately, most cases of cervical cancer are now curable but the clinical impact of curative radiation and inoperable residue still poses an important challenge in terms of patient health [2]. Over the past 30 years, imaging has focused on the classification of cervical carcinoma to differentiate patients better treated either by surgery or radiotherapy [3]. Response evaluation to radiation therapy is also a contemporary issue, specifically due to new MRI sequences that may be able to differentiate tumours from oedema [4].

Women are particularly vulnerable to disease in their reproductive years (the early teens to about 50 years of age). In addition to the risks of pregnancies, specific issues are of concern in the general population. For instance, the menstrual cycle can place a significant burden on patients suffering from endometriosis, and there may also be knock-on effects on reproduction in terms of poorer prospects for pregnancy [5]. Although imaging has been widely used, particularly MRI due to its ability to non-invasively overview the disease status of the pelvis [6], to the best of our knowledge no literature has yet described the use of MRI in the screening of young adults for endometriosis.

Although uncommon, specific hepatocellular tumours are almost exclusively present in females, mainly during their reproductive years [7]. Hepatocellular adenoma's may require urgent intervention since they are prone to bleeding. In rare cases they may also deteriorate to a malignant form, referred to as hepatocellular carcinoma.

Two specific topics in women's abdominal health will be researched in this thesis: firstly, pelvic MRI, particularly involving cervical carcinoma and endometriosis, and secondly, the MRI of hepatocellular adenoma.



## Part 1

### Cervical carcinoma

Cervical cancer represents a major health burden, with for instance around 12,000 newly diagnosed patients in the United States each year (US Cancer Statistics Working Group 2016). Until recently, cervical cancer staging was performed with examination under anaesthesia (EUA) according the Fédération Internationale de Gynécologie et d'Obstétrique (FIGO) classification [8]. The main focus of this classification is to identify operable patients (FIGO < IIB without FIGO IB2) versus those patients better treated with radiation therapy, with or without concurrent chemotherapy (FIGO > IIA and FIGO IB2) or radiotherapy [9].

Since its inception this classification has undergone various revisions, but only in 2009 was MRI definitely accepted as an adjunct and even as an alternative to EUA [9]. Strikingly, MR imaging was already proposed and evaluated for the purpose of the detection of parametrial invasion in 1986 [10]. Possible explanations for why MR imaging has only recently achieved acceptance include the heterogeneity of early results and the lack of availability of MRI in many countries (mainly in the developing world, where the majority of cervical cancer patients are actually found). To date, no systematic review has compared the value of clinical examination with MRI. A review is sorely needed not only because MRI now plays such an important role in diagnostics, but also because MRI is now widely used in radiotherapy planning [11-15] and more recently, in MR imaging-based hyperthermia [14;16]. Greater clarity regarding the value of MRI is therefore urgently needed.

At our institution, the work-up of a patient with proven cervical carcinoma was, until recently, mainly based on EUA, tumor markers and CT. EUA is a relatively invasive and expensive procedure that requires operating room time, anesthesia equipment, (often) three physicians (a gynecologist, radiotherapist and anesthesiologist), and supporting personnel. In some cases, the examination can be completed with a cystoscopy or rectoscopy [9]. By contrast, MRI has the potential to evaluate the cervix less invasively and at lower expense.

The goal of our study in **chapter 2** was to systematically review the literature on the diagnostic performance of EUA and MRI in detecting parametrial invasion and advanced stage disease in patients with primary cervical carcinoma, using surgico-pathological results as a reference standard.

MRI has also been proposed as an adjunct to clinical examination in the response evaluation of cervical carcinoma after radiotherapy. However, the main problem seemed to be that conventional MRI produced a large number of false positives, mainly due to post-therapeutic edema. Edema leads to a bright cervix on T2-weighting, leading to difficulties in distinguishing local tumor residue [17].

Diffusion MRI has recently been proposed as a useful alternative in such situations, allowing edema to be differentiated from local tumor burden [4]. However, these data were derived from small patient samples and, in most studies, using reference standards of poor quality. Therefore, no reliable evidence exists to date to support use of MRI plus DWI, alongside clinical response evaluation, in the response evaluation of cervical carcinoma after radiotherapy. Should MRI plus DWI be added to the standard protocol? This might lead to an increase in false positives, with ensuing unwelcome anxiety for patients.

In **chapter 3**, we prospectively studied the value of MRI plus DWI compared to clinical examination, in order to determine the added value of MRI plus DWI and its potential as an alternative to clinical examination. In this study, we used a reliable reference standard consisting of pathologic evidence and follow-up of at least one year.

## Endometriosis

Endometriosis is another disease in which MRI could play an important role in staging, and possibly also in detection of the disease. Studies of the utility of MRI in endometriosis have so far mainly focused on the preoperative setting in order to map the locations of deep endometriosis in bladder and rectum [18;19]. This mapping allows the gynecologist to plan the operation, should it be performed. However, the use of MRI to diagnose endometriosis, irrespective of the presence and extent of deep or superficial endometriosis, has not been fully addressed until now. Development of a non-invasive test for endometriosis has long been an important priority in endometriosis research because the delay that precedes an accurate diagnosis of endometriosis can be substantial and secondly, early detection of the disease using a non-invasive approach such as MRI might prevent further deterioration in the pelvis (which is accompanied by secondary pain and subfertility) [20]. The purpose of our study, explained in **chapter 4**, was to explore whether an optimized MRI protocol using 3.0-Tesla is sufficiently sensitive to detect the disease in a clinical setting of cyclic pain or subfertility, compared to laparoscopic exploration.

## Part 2

### Hepatocellular adenoma

Hepatocellular adenomas (HCA) are uncommon, primarily benign tumors of the liver and are almost exclusively apparent in females [21]. HCA typically affects young women of childbearing age with a history of contraception [21]. Although rare, tumors in males are more prone to malignant degeneration to hepatocellular carcinoma thus require closer follow-up [22].

Although benign in nature, prompt diagnosis and differentiation from other tumors is very important. Once the diagnosis of HCA is made, patients should be informed that they have a small but increased risk of acute tumor bleeding and that there is risk of malignant

degeneration, particularly in larger tumors [23]. Since these tumor characteristics are influenced by female hormones, hormonal contraception should be avoided and pregnancy should be discouraged [23;24]. This has a major impact on the lives of these women, and therefore accurate diagnosis is of primary importance.

Diagnosis can be established by biopsy, but this invasive technique is not preferred as there is a risk of serious bleeding during the procedure [24]. Generally, the diagnosis is established non-invasively with MR imaging or contrast sonography. The latter technique is currently under investigation [25].

### **New subclassification of HCA (Bordeaux classification)**

A molecular classification of HCA has been introduced by the Bordeaux group [26]. They described two genetic alterations occurring in these lesions, namely the inactivation of hepatocyte nuclear factor 1 alpha (HNF1 $\alpha$ ) and the activating mutation of  $\beta$ -catenin. A third group was initially termed telangiectatic focal nodular hyperplasia (telangiectatic FNH), but is now referred to as HCA since its behaviour more closely resembles HCA than FNH [27].

In 2007, the same group identified several (immunohistochemical) histopathological markers that are highly sensitive and specific in the differentiation of the subtypes [26].

There are currently four immunohistochemically recognized subtypes of HCA, classified respectively as inflammatory HCAs (40–50%, IHCA), *HNF1A*-mutated HCA (30–40%, H-HCA),  $\beta$ -catenin activated HCAs (10–15% b-HCA), and unclassified HCAs (10–25%, U-HCA). They all show characteristic pathological findings and there seems to be a correlation between some subtypes and clinical behaviour.

While immunohistochemical staining of LFABP,  $\beta$ -catenin, glutamine synthetase, serum amyloid A, and C-reactive protein has proven to be very effective in differentiating between the four subtypes of HCA, it is also useful in the differentiation between HCA and FNH [28].

### **Differentiation HCA from FNH**

FNH is an asymptomatic benign tumour with normal liver enzyme levels and tumour markers that is mostly found incidentally using imaging [29]. FNH has a benign behavior and, once diagnosed, usually requires no follow-up. The major challenge in the diagnosis of FNH is differentiation from HCA. Both FNH and IH-HCA are characterized by positive GS-staining on histochemical analysis [30].

A map-like pattern seems to be exclusive to FNHs and is not present in HCAs, making the differentiation in most cases straightforward [31]. In cases of small biopsies this can, however, be challenging. Furthermore, both FNH as IH-HCA present with internal bile ducts, in contrast to the other subtypes of HCAs [24]. Before the introduction of this typical presentation of

map-like GS staining, some IH-HCA was wrongly classified as FNHs due to internal bile duct proliferation. However, the clinical presentation of IH-HCA is clearly different from FNH, as IH-HCA shows a propensity for internal bleeding and malignant degeneration, neither of which are found in FNHs. A final problem in diagnosing FNHs is achieving differentiation from HCAs by imaging, since they both may have quite comparable features on imaging. Our institution and others have described some typical MRI features that can assist in the differentiation of these two entities [32;33]. However, in these studies the reference standard was based on classical histological findings without support from immunohistochemistry, raising the possibility of misclassification. Consequently, there is a need for reevaluation of typical FNH and HCA MRI findings based on advances in immunohistochemistry.

Hepatobiliary MRI contrast agents such as gadobenate dimeglumine and gadoxetate disodium are widely available, and they appear to be sufficiently discriminating to allow differentiation of HCAs from FNHs. This conclusion is mainly based on the fact that FNH excrete this contrast medium in the intra-lesional bile ducts, causing these lesions to appear bright in the hepatobiliary phase [34]. Differentiation from FNH appears to be relatively straightforward because HCAs are typically not bright in the hepatobiliary phase. Although this is the picture portrayed in current literature, once again the reference standard until recently was often unsupported by immunohistochemical staining, and may therefore have been subject to misinterpretation [34]. This particular difficulty arose because IH-HCA resembles FNH using classical staining without GS, an error partly attributable to the fact that these lesions harbor internal bile ducts, which was thought to be restricted to FNH [24]. Although not yet evaluated, it would be interesting to analyze the behavior of IH-HCA after injection of a hepatobiliary contrast agent. Furthermore, although both contrast agents are thought to be interchangeable, this has never been proven. The best method to address this question would be to scan HCA or FNH patients with both contrast agents, and to look for differences in the hepatobiliary phase.

Once the diagnosis of HCA is made, treatment options are diverse and possibly complex. Until recently there was little international agreement on how to treat and follow-up these lesions, and although a wait-and-see policy is often proposed, this cannot always be justified [35].

One of the aims of this thesis was to clarify the MR imaging findings typical of HCAs, particularly in comparison to FNH, and to search for MR findings that could discriminate the different subtypes of HCAs, as is elaborated in **chapter 5**. Naturally, this research used immunohistochemical staining as a reference standard. The usefulness of hepatobiliary contrast agents in the differentiation of HCA from FNH will be further explored in **chapter 6**, **chapter 7**, **chapter 8**, and **chapter 9**. Furthermore, in **chapter 10** a proposal will be formulated how to deal with these lesions in clinical practice.

## References

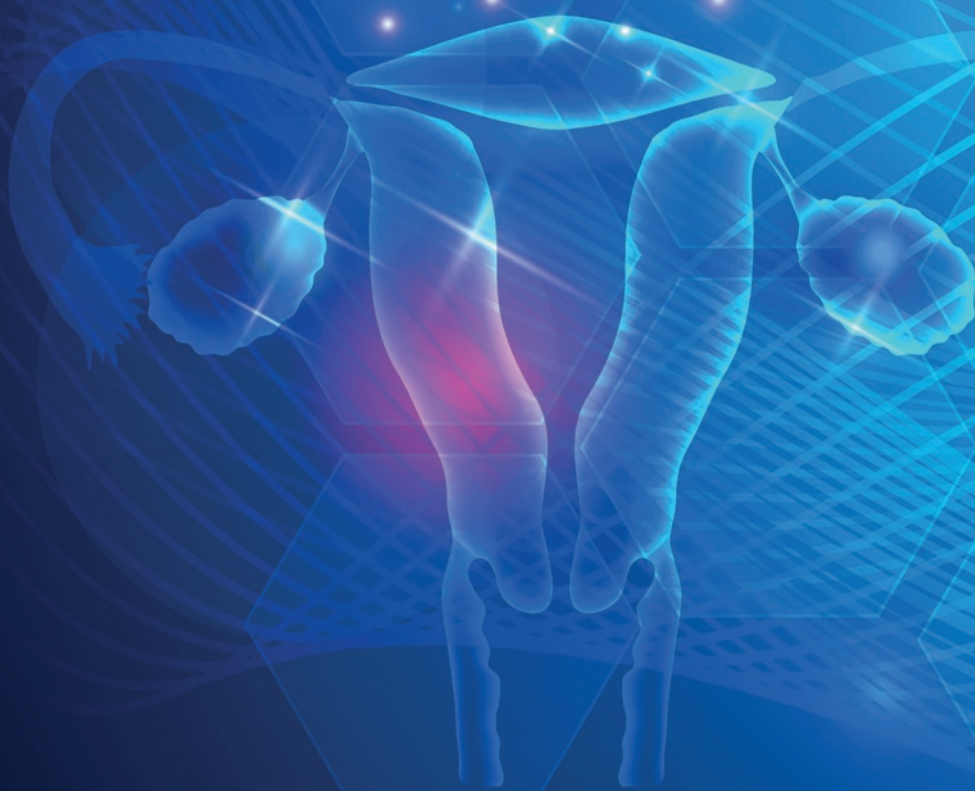
1. Papanicolaou GN (1942) A New Procedure for Staining Vaginal Smears. *Science* 95:438-439
2. Kim JY, Byun SJ, Kim YS, Nam JH (2017) Disease courses in patients with residual tumor following concurrent chemoradiotherapy for locally advanced cervical cancer. *Gynecol Oncol* 144:34-39
3. Hricak H, Lacey CG, Sandles LG, Chang YC, Winkler ML, Stern JL (1988) Invasive cervical carcinoma: comparison of MR imaging and surgical findings. *Radiology* 166:623-631
4. Schreuder SM, Lensing R, Stoker J, Bipat S (2015) Monitoring treatment response in patients undergoing chemoradiotherapy for locally advanced uterine cervical cancer by additional diffusion-weighted imaging: A systematic review. *J Magn Reson Imaging* 42:572-594
5. Kennedy S, Bergqvist A, Chapron C et al (2005) ESHRE guideline for the diagnosis and treatment of endometriosis. *Hum Reprod* 20:2698-2704
6. Bazot M, Bharwani N, Huchon C et al (2016) European society of urogenital radiology (ESUR) guidelines: MR imaging of pelvic endometriosis. *Eur Radiol*. 10.1007/s00330-016-4673-z
7. 10.1007/s00330-016-4673-z [pii]
8. Vijay A, Elaffandi A, Khalaf H (2015) Hepatocellular adenoma: An update. *World J Hepatol* 7:2603-2609
9. Pecorelli S, Benedet JL, Creasman WT, Shepherd JH (1999) FIGO staging of gynecologic cancer. 1994-1997 FIGO Committee on Gynecologic Oncology. International Federation of Gynecology and Obstetrics. *Int J Gynaecol Obstet* 64:5-10
10. Pecorelli S, Zigliani L, Odicino F (2009) Revised FIGO staging for carcinoma of the cervix. *Int J Gynaecol Obstet* 105:107-108
11. Togashi K, Nishimura K, Itoh K et al (1986) Uterine cervical cancer: assessment with high-field MR imaging. *Radiology* 160:431-435
12. Dimopoulos JC, Schirl G, Baldinger A, Helbich TH, Potter R (2009) MRI assessment of cervical cancer for adaptive radiotherapy. *Strahlenther Onkol* 185:282-287
13. Gong QY, Tan LT, Romaniuk CS, Jones B, Brunt JN, Roberts N (1999) Determination of tumour regression rates during radiotherapy for cervical carcinoma by serial MRI: comparison of two measurement techniques and examination of intraobserver and interobserver variability. *Br J Radiol* 72:62-72
14. Conibear J, Lowe G, Hoskin PJ (2012) High-precision MRI-guided adaptive brachytherapy for cervical carcinoma. *Int J Hyperthermia* 28:501-508
15. Staruch RM, Hynynen K, Chopra R (2015) Hyperthermia-mediated doxorubicin release from thermosensitive liposomes using MR-HIFU: Therapeutic effect in rabbit Vx2 tumours. *Int J Hyperthermia*. 10.3109/02656736.2014.992483:1-16
16. Craciunescu OI, Thrall DE, Vujaskovic Z, Dewhirst MW (2010) Magnetic resonance imaging: a potential tool in assessing the addition of hyperthermia to neoadjuvant therapy in patients with locally advanced breast cancer. *Int J Hyperthermia* 26:625-637
17. Soares PI, Ferreira IM, Igreja RA, Novo CM, Borges JP (2012) Application of hyperthermia for cancer treatment: recent patents review. *Recent Pat Anticancer Drug Discov* 7:64-73
18. Vincens E, Balleyguier C, Rey A et al (2008) Accuracy of magnetic resonance imaging in predicting residual disease in patients treated for stage IB2/II cervical carcinoma with chemoradiation therapy: correlation of radiologic findings with surgicopathologic results. *Cancer* 113:2158-2165
19. Bazot M, Darai E (2005) Sonography and MR imaging for the assessment of deep pelvic endometriosis. *J Minim Invasive Gynecol* 12:178-185; quiz 177, 186
20. Bazot M, Darai E, Hourani R et al (2004) Deep pelvic endometriosis: MR imaging for diagnosis and prediction of extension of disease. *Radiology* 232:379-389
21. D'Hooghe TM, Mihalji AM, Simsa P et al (2006) Why we need a noninvasive diagnostic test for minimal to mild endometriosis with a high sensitivity. *Gynecol Obstet Invest* 62:136-138

22. Rabe T, Feldmann K, Grunwald K, Runnebaum B (1994) Liver tumours in women on oral contraceptives. *Lancet* 344:1568-1569
23. Goudard Y, Rouquie D, Bertocchi C et al (2010) [Malignant transformation of hepatocellular adenoma in men]. *Gastroenterol Clin Biol* 34:168-170
24. van Aalten SM, Witjes CD, de Man RA, IJzermans JN, Terkivatan T (2012) Can a decision-making model be justified in the management of hepatocellular adenoma? *Liver Int* 32:28-37
25. Thomeer MG, ME EB, de Lussanet Q et al (2014) Genotype-phenotype correlations in hepatocellular adenoma: an update of MRI findings. *Diagn Interv Radiol* 20:193-199
26. Laumonier H, Cailliez H, Balabaud C et al (2012) Role of contrast-enhanced sonography in differentiation of subtypes of hepatocellular adenoma: correlation with MRI findings. *AJR Am J Roentgenol* 199:341-348
27. Bioulac-Sage P, Blanc JF, Rebouissou S, Balabaud C, Zucman-Rossi J (2007) Genotype phenotype classification of hepatocellular adenoma. *World J Gastroenterol* 13:2649-2654
28. Paradis V, Benzekri A, Dargere D et al (2004) Telangiectatic focal nodular hyperplasia: a variant of hepatocellular adenoma. *Gastroenterology* 126:1323-1329
29. Bioulac-Sage P, Cubel G, Taouji S et al (2012) Immunohistochemical markers on needle biopsies are helpful for the diagnosis of focal nodular hyperplasia and hepatocellular adenoma subtypes. *Am J Surg Pathol* 36:1691-1699
30. Hussain SM, Terkivatan T, Zondervan PE et al (2004) Focal nodular hyperplasia: findings at state-of-the-art MR imaging, US, CT, and pathologic analysis. *Radiographics* 24:3-17; discussion 18-19
31. van Aalten SM, Verheij J, Terkivatan T, Dwarkasing RS, de Man RA, IJzermans JN (2011) Validation of a liver adenoma classification system in a tertiary referral centre: implications for clinical practice. *J Hepatol* 55:120-125
32. Bioulac-Sage P, Laumonier H, Laurent C, Zucman-Rossi J, Balabaud C (2008) Hepatocellular adenoma: what is new in 2008. *Hepatol Int* 2:316-321
33. van den Bos IC, Hussain SM, de Man RA, Zondervan PE, IJzermans JN, Krestin GP (2008) Primary hepatocellular lesions: imaging findings on state-of-the-art magnetic resonance imaging, with pathologic correlation. *Curr Probl Diagn Radiol* 37:104-114
34. Grazioli L, Morana G, Kirchin MA, Schneider G (2005) Accurate differentiation of focal nodular hyperplasia from hepatic adenoma at gadobenate dimeglumine-enhanced MR imaging: prospective study. *Radiology* 236:166-177
35. Grazioli L, Bondioni MP, Haradome H et al (2012) Hepatocellular adenoma and focal nodular hyperplasia: value of gadoteric acid-enhanced MR imaging in differential diagnosis. *Radiology* 262:520-529
36. Terkivatan T, Hussain SM, De Man RA, IJzermans JN (2006) Diagnosis and treatment of benign focal liver lesions. *Scand J Gastroenterol Suppl.* M172651155786T32 [pii]



# Part I

## Gynaecological imaging





# Chapter 2

## **Clinical examination versus Magnetic Resonance Imaging for staging of cervical carcinoma: systematic review and meta-analysis**

Maarten G Thomeer

Cees Gerestein

Sandra Spronk

Helena Van Doorn

Myriam Hunink

## Abstract

**Objectives:** To review the literature on the diagnostic performance of clinical examination and Magnetic Resonance Imaging (MRI) in detecting parametrial invasion and advanced stage disease (FIGO stage  $\geq$  IIB) in patients with cervical carcinoma.

**Methods:** Reports of studies were searched using the MEDLINE, EMBASE and Cochrane databases. Two observers reported on data relevant for analysis and methodological quality using the QUADAS scoring system. Publication bias was analysed using Deeks funnel plots. Covariates were added to the model to study the influence on the summary results of the technical and methodological aspects of the clinical examination and MRI.

**Results:** In total, 3254 patients were included. Partial verification bias, was often encountered. Pooled sensitivity was 40% (95% CI:25–58) for the evaluation of parametrial invasion with clinical examination and 84% (95% CI:76–90) with MRI, 53% (95% CI:41–66) for the evaluation of advanced disease with clinical examination, and 79% (95% CI: 64–89) with MRI. Pooled specificities were comparable between clinical examination and MRI. Different technical aspects of MRI influenced the summary results.

**Conclusions:** MRI is significantly better than clinical examination in ruling out parametrial invasion and advanced disease in patients with cervical carcinoma.

### *Keywords*

MRI, cervix, parametrial, FIGO, meta-analysis

### *Key points*

1. MRI has a higher sensitivity than clinical examination for staging cervical carcinoma.
2. Clinical examination and MRI have comparably high specificity for staging cervical carcinoma.
3. Quality of clinical examination studies was lower than that of MRI studies.
4. The use of newer MRI techniques influences positively the summary results.
5. Anaesthesia during clinical examination influences positively the summary results.

## Introduction

Carcinoma of the uterine cervix is the second most common malignancy in women, with an estimated worldwide incidence of 530,000 cases per year and is characterised by a high case fatality rate (52%) [1]. The Federation Internationale de Gynecologie et d'Obstetrique (FIGO) staging system is used for the clinical staging of cervical carcinoma and is mainly based on tumour diameter and local extent [1]. FIGO staging is currently determined clinically by inspection and rectovaginal palpation, with or without anaesthesia, to allow evaluation of advanced disease. According to the FIGO Committee on Gynecologic Oncology further staging may be performed with additional examinations such as cystoscopy and proctoscopy, but CT and MRI are only recommended, but not mandatory [2].

The utility of CT for cervical staging has been shown to be inferior to that of Magnetic Resonance Imaging (MRI) [3]. MRI has been extensively evaluated and compared with standard CE, but clear evidence that MRI is a diagnostic improvement in comparison to CE, or should even replace CE, is still lacking. This is probably the reason why MRI was not used by up to 30% of the responders in a very recent survey among the members of the Gynaecologic Cancer Intergroup [4]. Finally, the quality of studies analysing these two diagnostic procedures have never been systematically compared using standardised quality assessment scores.

The purpose was to review the literature on the diagnostic performance of CE and MRI in detecting parametrial invasion and advanced stage disease (FIGO stage  $\geq$  IIB) in patients with cervical carcinoma, using surgico-pathological results as a reference standard.

## Materials and methods

### Protocol and registration

Methods of analysis and inclusion criteria were specified in advance and documented in a protocol (Dutch Trial Registry NTR2896).

### Eligibility criteria and search terms

Articles were identified by searches using the PICO criteria (see below), with restriction of language (only studies in the English language were included) and publication status (no unpublished data or abstracts were included).

- Types of participants (P): patients of any age with cancer (e.g. cancer, carcinoma, malignancy or neoplasm) of the cervix.
- Types of intervention (I): studies where clinical investigation (e.g. gynaecological examination, clinical examination, examination under anaesthesia, or EUA) and/or magnetic resonance imaging (e.g. MRI or MR or NMR) were performed.

- Types of comparison (C): studies where staging (staging, stage, FIGO or TNM) was performed.
- Types of Outcome (O): sensitivity and specificity.

A complete list of the search strategies can be found in the appendix.

### **Information resources**

Studies were identified through searches of electronic databases (MEDLINE, Cochrane database and EMBASE) and the reference lists of the full text papers included. The search was applied without a starting date, with an update until the end of December 2011.

### **Study selection and data collection process**

Two board certified investigators (MT; radiologist and CG; gynaecologist) performed a structured and independent search through the databases for original articles. In cases of conflicting views as to whether a study should be included, a consensus decision was reached. If no agreement could be reached, a third investigator decided on final inclusion (HD; gynaecologist). The search was primarily based on the title, subsequently on the abstract and finally on the full article.

Exclusion criteria were as follows: the article was not written in English, the article was not about adenocarcinoma or squamous or adeno-squamous carcinoma, the study was not about diagnostic tests, the patients included in the study had already been treated for cervical carcinoma, the study included less than ten patients, the patient group included pregnant women, the study did not assess MRI or clinical examination as a diagnostic tool, inability to obtain original numbers of true-positives, false-positives, true-negatives and false-negatives, or overlap of patient data in different studies.

### **Data items**

From the full articles included, several overall data items were extracted independently by two authors (MT and CG) as follows:

*the year of publication, the number of patients included, the prevalence of parametrial invasion and advanced disease, the type of study (uni- or multicentre), the design (prospective or retrospective), the consecutive nature of the data, the indication of inoperability, the reason for surgery (CE or MRI), the specifications of the CE (e.g. which investigator(s), rectovaginal examination, anaesthesia, cystoscopy and rectoscopy), the specifications of the MRI (e.g. spin echo [SE] or fast spin echo [FSE] sequences, the use of contrast medium, the use of a bowel relaxant, the type of coil and the field strength).*

Finally, from the selected papers, the numbers of true-positives, true-negatives, false-positives and false-negatives were extracted for CE and MRI. This was performed for parametrial invasion and advanced disease (also mentioning FIGO IIB or higher including IB2 as advanced



disease if reported). If studies reported different MRI techniques in the same patients, only the highest sensitivity was extracted.

## Quality assessment of individual studies

In order to assess the quality of the diagnostic studies, the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tool was used [5]. This consists of a set of 14 items that encompass the different biases typically encountered in diagnostic studies. The QUADAS score was based on consensus reading of two authors (MT and CG) independently (Figure 1). Discrepancies were resolved by consensus.

We added a fifteenth bias, a test review bias, which can occur when two different diagnostic tools are used in the same study. In these cases the knowledge of the outcome of one examination can influence the evaluation of the other.

1. Was the spectrum of patients representative of the patients who will receive the test in practice?
2. Were selection criteria clearly described?
3. Is the reference standard likely to classify the target condition correctly?
4. Is the period between the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests?
5. Did the whole sample or a random selection of the sample receive verification using a reference standard of diagnosis?
6. Did patients receive the same reference standard regardless of the index test result?
7. Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?
8. Was the execution of the index test described in sufficient detail to permit replication of the test?
9. Was the execution of the reference standard described in sufficient detail to permit its replication?
10. Were the index test results interpreted without knowledge of the results of the reference standard?
11. Were the reference standard results interpreted without knowledge of the results of the index test?
12. Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?
13. Were uninterpretable/intermediate test results reported?
14. Were withdrawals from the study explained?

**Figure 1.** QUADAS score system used for diagnostic studies and based on 14 questions. QUADAS = Quality Assessment of Diagnostic Accuracy Studies [5]

## Data synthesis and analysis

A bivariate random effects regression model was used for data analysis [6]. This model assumes a binomial distribution of the within-study variability (variability between sensitivity and specificity within a study). This model furthermore assumes correlated and normally distributed random effects between studies. The degree of correlation between the logit sensitivity and logit specificity corresponds to the inverse relation between sensitivity and specificity when the positivity criterion is varied.

The data of each study were summarised in forest plots to obtain summary estimates of sensitivity and specificity, with 95% confidence intervals for both CE and MRI. This was performed for parametrial invasion and advanced disease separately. Positive likelihood (LR+), negative likelihood (LR-) and the diagnostic odds ratio (DOR) were also calculated.

Heterogeneity of sensitivity and specificity was assessed using the  $I^2$  statistic [7], which is a derivative of the Cochran Q statistic [8]. Whereas the Q statistic informs us about the presence

of heterogeneity, the  $I^2$  statistic quantifies the degree of heterogeneity. The larger the  $I^2$  value the more heterogeneity exists across the studies. Both the Cochran Q and the  $I^2$  statistic display a low power for the detection of heterogeneity when the number of studies is small and a high power for detection when the number of studies is large.

Additionally, univariate meta-regression analyses were performed to explore the effect of the co-variables for both CE and MRI. Continuous variables explored were the year of publication and the prevalence of disease (parametrial invasion or detection of advanced disease). Dichotomous variables explored were the design of the study (prospective versus retrospective) and the absence of verification bias (all patients were operated versus not all patients were operated). Specifically for CE, the influence of anaesthesia on the summary results was analysed. For MRI, the following co-variables were also explored: the use of FSE sequence versus SE sequence, the use of a high field magnetic strength (1.5 Tesla or higher) versus a low field magnetic strength (lower than 1.5 Tesla), the use of an additional coil (phased array and endocoil) versus standard body coil, the use of contrast medium, and the slice thickness (1–5 mm versus larger than 5 mm).

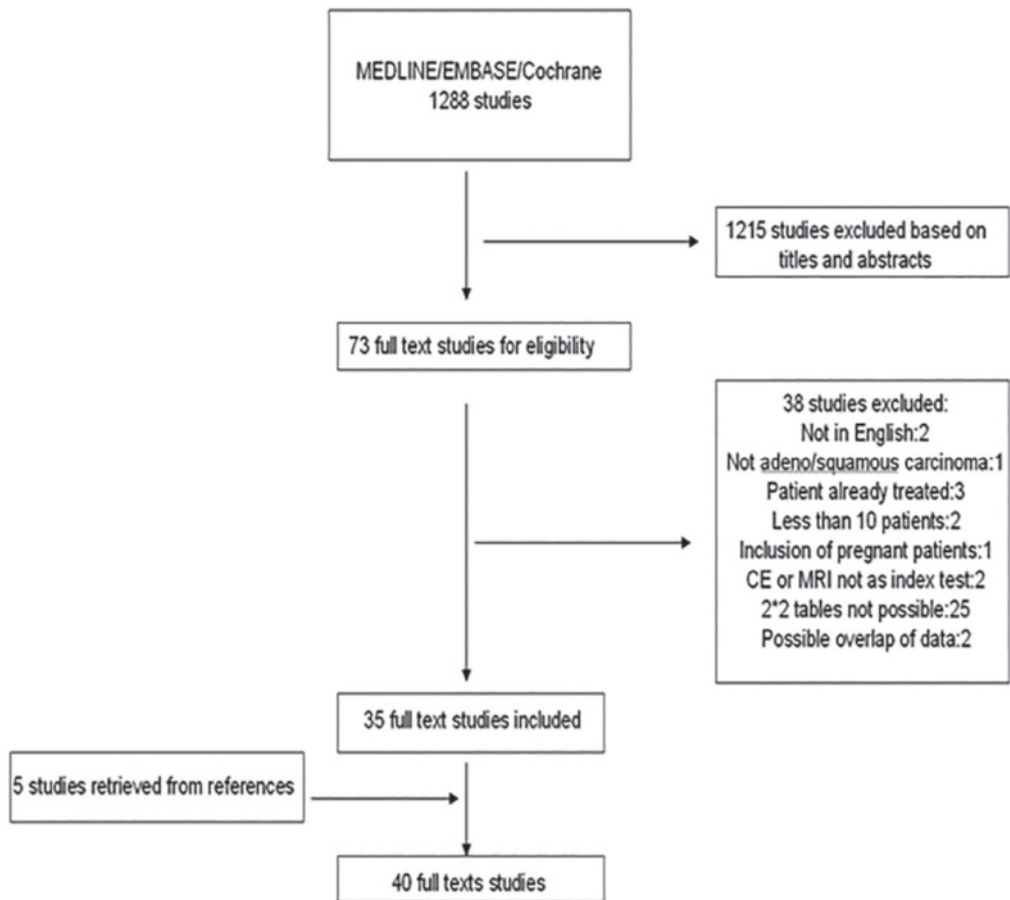
A variant of the funnel plot was used, the Deeks plot [9], for estimation of publication bias. The Deeks plot is preferred over the standard Egger and Smith [10] or Macaskill et al [11] plots because of its proven higher power especially when DORs are heterogeneous. The Deeks plot is used to examine the asymmetry of the DORs versus the effective sample sizes. A significant regression coefficient indicates an association between sample size and the DORs, i.e. the overall performance of small studies differs from that of larger studies. In that case, there can be an under- or over-reporting in smaller studies, referred to as publication bias.

A *P* value of 0.05 or less was considered statistically significant. Calculation and analyses were performed using Microsoft Excel 2010 (Microsoft, Seattle, WA, USA) and the MIDAS module for STATA, version 12 (StataCorp, College Station, TX, USA).

## Results

Based on the PICO criteria, 1288 unique studies were selected from MEDLINE, EMBASE and the Cochrane database (Figure 2). 1215 studies were excluded based on titles or abstracts.

From the included studies 38 were excluded because of: the article was not written in English (two), other tumors than adeno/squamous subtypes were included (one), patients were already treated (three), only less than 10 patients were included (two), pregnant women were included (one), CE or MRI was not used as index test (two), two by two tables were not possible to produce (25), and possible overlap of data was possible (two). Five additional studies were retrieved from reference lists [12-16].



**Figure 2.** Flow diagram of study selection by two authors (radiologist and gynaecologist). CE = clinical examination

In total, 3254 patients were included in 40 papers [4;12-50]. Twenty-one studies reported on MRI only, 4 on CE only and 15 on both tests. Concerning the analysis of parametrial invasion, some authors divided the parametrial invasion into left and right (nine studies) [12;18;21;36;37;39;41;44;47], but most authors used the term parametrial invasion as invasion on at least one side (21 studies) [4;13;14;19;20;24;28;30;31;33-35;38;40;42;43;45;46;48-50].

Tables 1 and 2 summarise the characteristics of CE and MRI. The median prevalence of parametrial invasion was 13% for CE (range: 5–33%) and 22% for MRI (range: 5–73%). The median prevalence of advanced disease was 22% for CE (range: 5–50%) and 23% for MRI (range: 5–85%).

Most studies were prospective in design (28/40; 70%) and consecutive in patient recruitment (minimally: 34/40; 85%). In most of the papers, patients with FIGO stage IIb or more were not

operated (14 papers). Among the latter studies, two also excluded patients from operation when the tumour was larger than 4 cm (FIGO stage Ib2). In five papers, mostly from the East, surgery was attempted as late as FIGO stage IIIa. The seven studies in which all patients were operated showed early dates (all pre-1990).

Details of the examination were often not reported in studies involving CE, with the exception of the three largest studies (62% of the patients) where most details were described.

For the studies involving MRI, details on the technique were described in more detail in the almost all of the studies (35/36; 97%).

Figures 3 and 4 summarise the results of the methodological quality assessment of the articles included based on the CE and MRI studies. The quality of the CE studies was lower, in general, in comparison to the MRI studies. Partial verification bias, but also differential verification bias was often present with both tests. The reference test was not blinded in most studies to the results of the CE. Finally, in most of the studies it was not evident whether the results of MRI and CE data influenced each other.

Sensitivities and specificities in the evaluation of parametrial invasion and advanced disease are shown in Table 3 and were found to be heterogeneous (i.e. high  $I^2$ ) for both CE and MRI, and this justifies our choice of using a random effects model, which results in a larger confidence interval for all sensitivities and specificities (Figures 5A-D8).

Larger differences in sensitivities for analysis of parametrial invasion and advanced disease were apparent between CE studies than in MRI studies, varying between 0 and 100%. However, in general, the heterogeneity factor  $I^2$  was comparable for both tests. Three studies had a sensitivity of 0% and specificity of 100% concerning advanced disease, which was mainly because most patients were not operated when advanced disease was clearly present during CE. Two MRI studies showed sensitivity below 50% for analysis of parametrial invasion, which was possibly due to the use of an older technique (SE with body coil and larger slice thicknesses). Concerning the evaluation of advanced disease, one study showed a sensitivity of 0% and specificity of 100%. This could be partly explained by the very low prevalence of disease in the operated group.

**Table 1.** Characteristics of all the studies included where clinical examination of the cervix was used as an index test

	Year	Number included	Prevalence parametrial invasion (%)	Prevalence advanced disease (%)	Multi-centre	Design	Executive	RVT	Anaesthesia	Cystoscopy	Rectoscopy	Consecutive	Minimal inoperable stage	Surgery based on
Averette et al [15]	1972	70	NA	40	–	P	NR	+	+	+	+	+	All stages operated	NA
Sundersanam et al [16]	1978	220	NA	40	–	P	NR	+	+	+	+	+	All stages operated	NA
Maggino et al [17]	1983	147	NA	23	–	R	NR	NR	NR	+	+	+	All stages operated	NA
Togashi et al [18]	1986	19	11	17	–	P	NR	+	NR	+	+/-	NR	All stages operated	NA
Hricack et al [19]	1988	57	33	33	–	R	G	+	NR	+/-	+/-	+	IIB	NR
Waggenspack et al [20]	1988	20	10	10	–	R	G	+	+	+	+	+	All stages operated	NA
Kim et al [21]	1990	30	20	30	–	P	G	NR	NR	NR	NR	+	NR	NR
Soeters et al [22]	1991	11	NA	27	–	R	1 G and 1 RT	+	NR	+	–	–	NR	NR
Yang et al [14]	1996	20	5	5	–	P	G	NR	+	NR	NR	+	NR	CE
Preidler et al [23]	1996	10	NA	50	–	P	2 G	NR	–	–	–	+	IIIA	NR
Kim et al [24]	1997	28	11	NA	–	P	NR	NR	NR	NR	NR	+	IIB	NR
Vierzen et al [25]	1998	26	NA	15	–	P	2 G and 1 RT	+	+	+	+	+	IIB	CE and MRI
Postema et al [26]	2000	91	NA	20	–	P	G	+	+	+/-	+/-	+	IIB	CE

Table 1. Continued

	Year	Number included	Prevalence parametrial invasion (%)	Prevalence advanced disease (%)	Multi-centre	Design	Executive	RVT	Anaesthesia	Cystoscopy	Rectoscopy	Consecutive	Minimal inoperable stage	Surgery based on
Hansen et al [27]	2000	61	NA	8	-	P	NR	+	+	+	+	+	IIB	NR
Wang et al [28]	2001	18	6	6	-	P	G	+	+	NR	NR	+	NR	NR
Hricack et al [29]	2005	165	NA	22	+	P	NR	+	+/-	+/-	+/-	+	IIB	NR
Chung et al [30]	2007	119	15	NA	-	R	NR	+	NR	+/-	+/-	+	IIIA	NR
Qin et al [13]	2009	818	10	10	+	R	2 G	+	-	+/-	+/-	+	IIIA	CE

NA = non-applicable; P = prospective; R = retrospective; NR = not reported; CE = clinical examination; MRI = magnetic resonance imaging; RVT = rectovaginal toucher; FIGO = Federation Internationale de Gynecologie et dObstetrique; G = gynaecologist; RT = radiotherapist



**Table 2.** Characteristics of all the studies included where MRI of the cervix was used as the index test

	Year	Number included	Prevalence parametrial invasion (%)	Prevalence advanced disease (%)	Multi-centre	Design	T2-weighted sequence	Intravenous Contrast medium	Bowel relaxant	Phased array/surface Coil	Field Strength 1.5 T or higher	Consecutive	Minimal inoperable FIGO stage	Surgery based on
Togashi et al [18]	1986	19	11	17	–	P	SE	–	–	–	+	NR	All stages operated	NA
Javitt et al [31]	1987	20	30	NA	–	P	SE	–	–	–	–	NR	All stages operated	NA
Hricack et al [19]	1988	57	33	33	–	R	SE	–	–	–	+/-	+	IIB	NR
Waggenspack et al [20]	1988	20	10	10	–	R	SE	–	–	–	+	+	All stages operated	NA
Togashi et al [32]	1989	66	NA	27	–	P	SE	–	–	–	+	+	IIB	NR
Janus et al [33]	1989	22	45	NA	–	P	SE	–	–	–	–	NR	All stages operated	NA
Greco et al [34]	1989	46	15	NA	–	P	SE	–	–	–	–	+	NR	NR
Kim et al [21]	1990	30	20	NA	–	P	SE	–	–	–	+	+	NR	NR
Soeters et al [22]	1991	11	NA	27	–	R	SE	–	–	–	–	–	NR	NR
Sironi et al [35]	1993	25	34	NA	–	P	SE	–	+	–	–	+	IIB	NR
Lien et al [12]	1993	169	5	NA	–	P	SE	–	+	–	+	+	IIB	CE
Kim et al [36]	1993	98	13	18	–	P	SE	–	–	–	–	+	IIB	CE/MRI
Kaji et al [37]	1994	21	14	NA	–	P	SE	+	+	-/+	+	+	IIB	NR

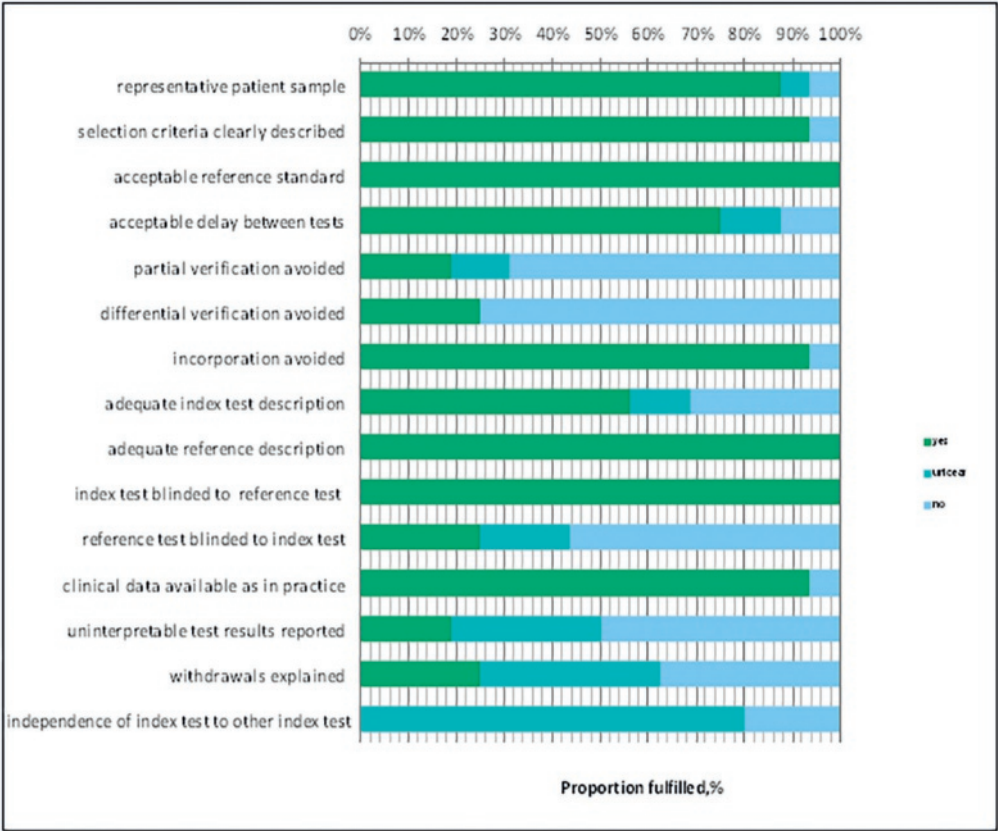
Table 2. Continued

	Year	Number included	Prevalence parametrial invasion (%)	Prevalence advanced disease (%)	Multi-centre	Design	T2-weighted sequence	Intravenous Contrast medium	Bowel relaxant	Phased array/surface Coil	Field Strength 1.5 T or higher	Consecutive inoperable FIGO stage on	Surgery based on
Subak et al [38]	1995	71	10	NA	–	R	FSE/SE	–	–	–/+	+	–	NR
Yang et al [14]	1996	20	5	5	–	P	SE	–	–	–	–	+	CE
Preidler et al [23]	1996	10	NA	50	–	P	FSE	–	–	+	+	+	NR
Hawighorst et al [39]	1996	26	73	85	–	P	FSE	+	–	–/+	+	+	NR
Kim et al [24]	1997	28	11	NA	–	P	FSE	–	–	+	+	+	NR
Scheidler et al [40]	1998	35	37	NA	–	P	FSE	+	–	+	+	+	CE
Hawighorst et al [41]	1998	32	70		–	P	FSE	+	–	+	+	+	NR
Vierzen et al [25]	1998	26	NA	15	–	P	FSE	+/–	+	+	+	+	CE/MRI
Yu et al [42]	1998	94	10	6	–	P	FSE/SE	–	+	+/–	+	+	NR
Ng et al [43]	1998	25	36	NA	–	R	NR	NR	NR	NR	NR	+	NR
Shiraiwa et al [44]	1998	50	19	NA	–	R	SE	–	–	–	+	+	NR
Postema et al [26]	2000	91	NA	20	–	P	FSE	–	+	–	+	+	CE
Hansen et al [27]	2000	61	NA	8	–	P	SE	+	–	–	–	+	NR
Wang et al [28]	2001	18	6	6	–	P	FSE	–	+	+	+	+	NR
Sheu et al [45]	2001	41	34	NA	–	P	FSE	+	–	+	+	+	NR
Oberoï et al [46]	2002	105	22	26	–	R	FSE	–	+	+	–	NR	NR

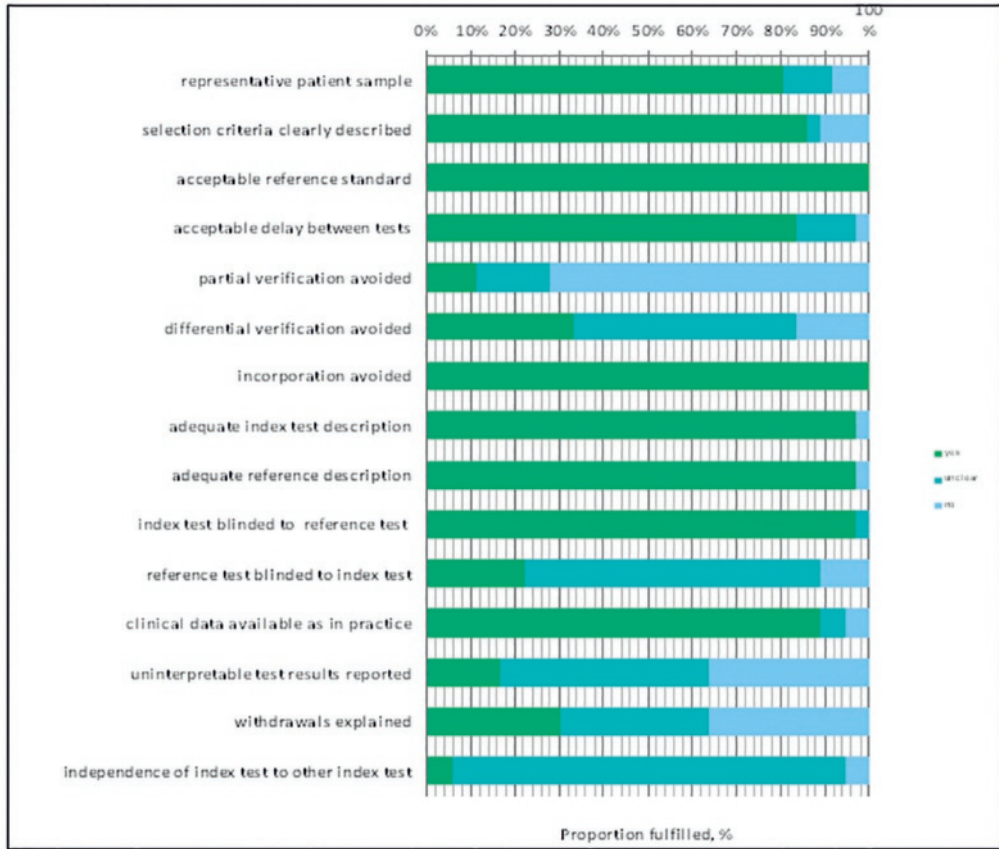
Table 2. Continued

	Year	Number included	Prevalence parametrial invasion(%)	Prevalence advanced disease(%)	Multi-centre	Design	T2-weighted sequence	Intravenous Contrast medium	Bowel relaxant	Phased array/surface Coil	Field Strength 1.5 T or higher	Consecutive	Minimal inoperable FIGO stage	Surgery based on
Choi et al [47]	2004	113	7	12	–	P	FSE	+	–	+	+	+	NR	NR
Desouza et al [48]	2006	119	13	NA	–	R	FSE	–	–	Endo	–/+	+	IIIA	NR
Chung et al [30]	2007	119	15	NA	–	R	FSE	–	–	+	+	+	IIIA	NR
Fischerova et al [49]	2008	95	6	NA	–	P	FSE	–	–	+	+	+	IIB	NR
Hori et al [50]	2009	31	13	NA	–	P	FSE	–	+	+	+	+	NR	NR
Hori et al [4]	2011	20	30	NA	–	R	FSE	–	+	+	+	+	NR	NR

NA: non-applicable; P = prospective; R = retrospective; NR: not reported; CE=clinical examination; MRI = magnetic resonance imaging; RVT: rectovaginal toucher; SE: spin echo sequence; FSE = fast spin echo sequence; T = Tesla; FIGO = Federation Internationale de Gynecologie et dObstetrique



**Figure 3.** Summary of quality of the clinical examination studies included, assessed using the QUADAS tool. The consensus judgement of the two readers is shown as cumulative percentages across the primary studies. QUADAS = Quality Assessment of Diagnostic Accuracy Studies



**Figure 4.** Summary of the quality of the MRI studies included, assessed using the QUADAS tool. The consensus judgement of the two readers is shown as cumulative percentages across the primary studies. QUADAS = Quality Assessment of Diagnostic Accuracy Studies; MRI: magnetic resonance imaging

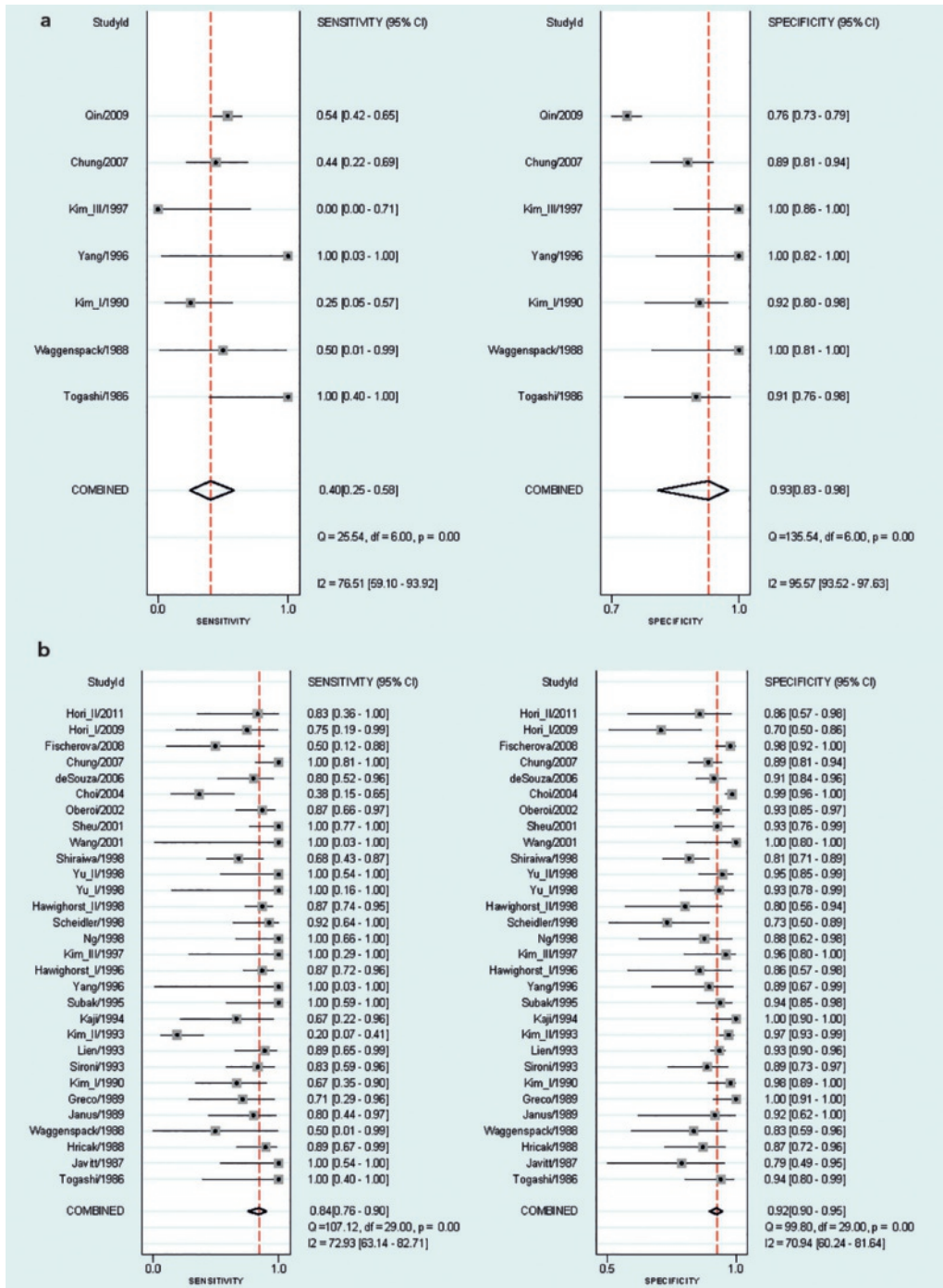
**Table 3.** Test characteristics of clinical examinations and MRI in the detection of parametrial invasion and advanced disease based on pooled data

TEST	Parametrial invasion		Advanced disease	
	Clinical examination	MRI	Clinical examination	MRI
Pooled Sensitivity	40% (95% CI:25–58)	84% (95% CI:76–90)	53% (95% CI:41–66)	79% (95% CI: 64–89)
Pooled Specificity	93% (95% CI:83–89)	92% (95% CI: 90–95)	97% (95% CI:91–99)	93% (95% CI: 88–96)
Positive LR	6.2 (95% CI:2.3–16.3)	11.10 (95% CI: 8.2–15.0)	19.3 (95% CI: 6.6–56.3)	11.2 (95% CI: 6.8–18.4)
Negative LR	0.64 (95% CI:0.5–0.8)	0.17 (95% CI: 0.1–0.3)	0.48 (95% CI: 0.4–0.6)	0.22 (95% CI: 0.1–0.4)
DOR	10 (95% CI:3–29)	65 (95% CI: 38–112)	40 (95% CI: 14–113)	50 (95% CI: 22–116)

LR = likelihood ratio; DOR = diagnostic odds ratio; CI = confidence interval

Different technical aspects of MRI influenced the summary results, both for parametrial invasion and the detection of advanced disease ( $P < 0.05$ ). The prevalence of disease in the studies had a significant effect on the summary results of MRI concerning parametrial invasion and on CE concerning the detection of advanced disease. The presence of verification bias had a negative effect on the summary results of MRI concerning parametrial invasion and advanced disease. The use of anaesthesia had a positive effect on the summary results of CE concerning the detection of advanced disease. The use of an FSE sequence, higher magnetic field, and an additional coil had a significant positive influence on the summary results of MRI for both parametrial invasion and the detection of advanced disease. The use of a slice thickness of 5 mm or smaller also had a significant positive effect on the evaluation of parametrial invasion with MRI. Other parameters were not significant or not interpretable because of the low number of studies included.

Publication bias, measured by Deeks funnel plot, was only significantly present in studies of parametrial invasion that used CE as an index test (Figures 6A-D) (Figures 9–12). Smaller studies (upper half) reported lower DORs in contrast to larger studies (lower half), revealing what was probably an underestimation of the results in the smaller studies.



**Figure 5 (a-d).** Forest plot for summary estimates of sensitivity and specificity with 95% confidence intervals for both clinical examination and MRI. If the  $I^2$  value is high, data have to be considered heterogeneous. Forest plots are measured for parametrial invasion using clinical examination (a) and MRI (b) as the index test. Forest plots are also measured for evaluation of advanced disease using clinical examination (c) and MRI (d) as the index test. StudyId = Study Identification



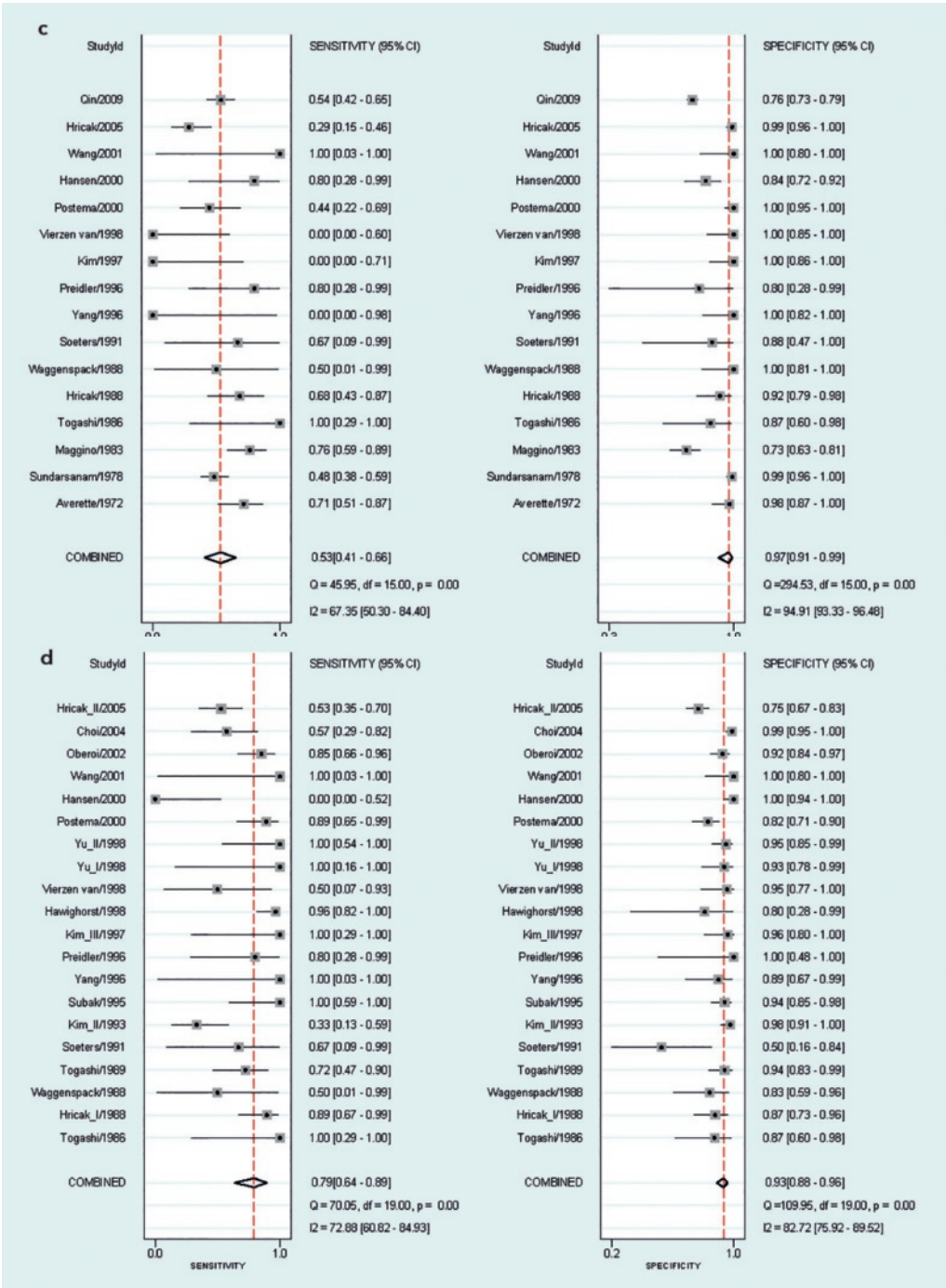
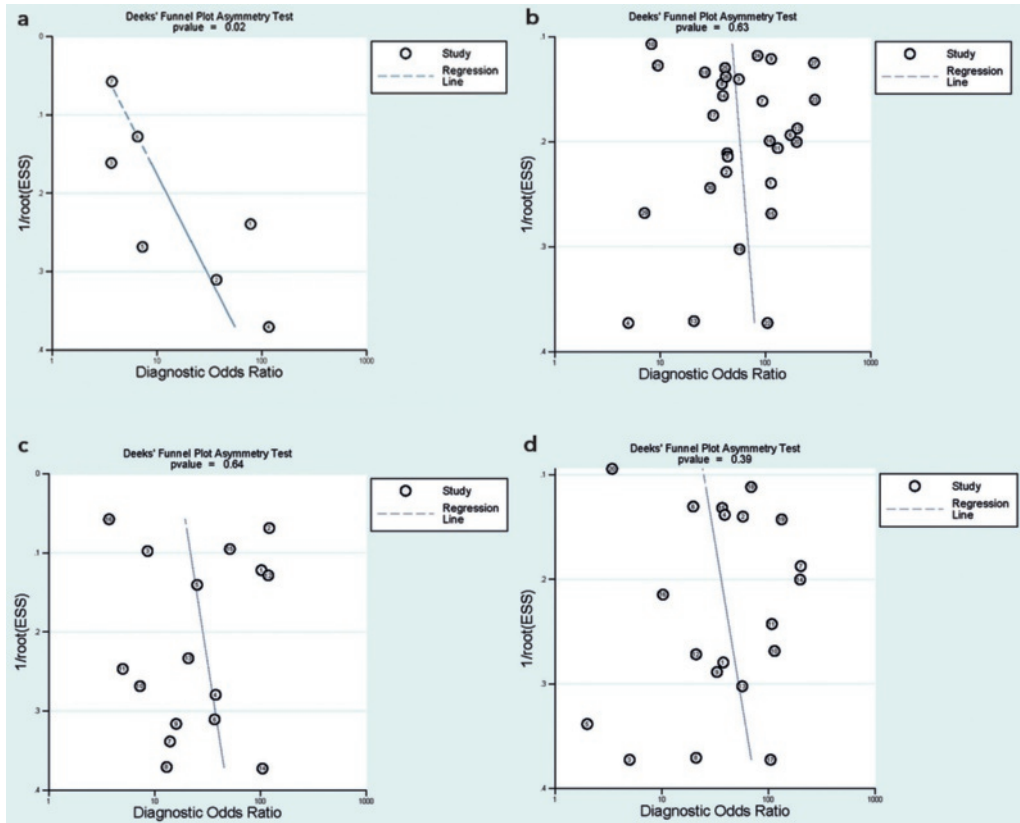


Figure 5 (a-d). Continued





**Figure 6 (a-d).** Deeks' funnel plot for the evaluation of asymmetry between the studies for the detection of parametrial invasion using clinical examination (a) and MRI (b) as the index test. C and d represent Deeks' funnel plot for the evaluation of advanced disease using clinical examination (c) and MRI (d) as the index test. Only in a was publication bias visible: smaller studies (upper half) reported lower diagnostic odds ratios (DORs) in contrast to larger studies (lower half), revealing what was probably an underestimation of the results in the smaller studies. ESS = effective sample size.  $P$  value < 0.05 means that there is asymmetry between the DORs of larger studies and smaller studies.

## Discussion

Our analysis of the available literature on CE and MRI revealed that the pooled sensitivities for the detection of parametrial invasion and advanced disease in cervical carcinoma are significantly higher for MRI than for CE and that the specificities are comparable and high.

The most important test characteristic in the evaluation of cervical cancer is sensitivity. Indeed, a test with a high sensitivity can rule out disease if the test is negative. The consequence of a low sensitivity, as is the case for CE, is that many patients will be operated initially, but will subsequently require postoperative (chemo)radiotherapy owing to unsuspected parametrial invasion or advanced disease. Multi-technique treatment increases morbidity significantly in comparison to treatment with only surgery or only radiotherapy. Side effects such as leg swelling (lymphatic obstruction), sexual dysfunction, urinary frequency, diarrhoea or constipation and bowel obstruction can occur more frequently in those patients who received postoperative radiotherapy [51-53]. Also, medical costs are higher when multi-technique treatment is needed.

A high specificity, as was found for both tests, is less relevant. A low specificity implies that when the test is positive, patients often receive unjustified radiotherapy instead of surgery. As the effectiveness of radiotherapy at these lower stages of cervical cancer is found to be nearly equivalent to surgery this seems less relevant at first [54;55]. However long-term morbidity is more pronounced when using radiotherapy as the main therapy choice.

Using a quality score to detect shortcomings in the methodology of the studies included, we discovered that partial verification bias was an important weakness in both CE and MRI studies. All patients underwent operations in early studies, but over the last few decades radical surgery for stage IIB and higher became obsolete and chemotherapy and/or radiotherapy is now advised [56;57]. Consequently, in several studies patients received no reference standard test once local invasion was detected, and therefore the presence of parametrial invasion and advanced disease documented with the reference standard test (surgery) was low in these studies.

The low prevalence of parametrial invasion and advanced disease in the studies retrieved is likely to be associated with a shift in case mix in the study population, which may have led to underestimation of the sensitivities of both tests. Patients with significant parametrial invasion or clear advanced disease will be easier to classify as true-positives, compared with less clear-cut cases. If these patients with evident disease were to have undergone the reference standard test, the rate of true-positives, and hence sensitivities, would have been much higher. Whether this partial verification bias affects the sensitivity of CE more than MRI is unclear, because in most studies it was not obvious whether CE and/or MRI determined the therapy

choice. For instance, if in a study CE determined whether surgery would be performed, the true-positives of CE included would be lower than those of MRI.

The specificities of both CE and MRI are very high and comparable to each other. As expected, the false-positive rate ( $= 1 - \text{specificity}$ ) was very low owing to false-positives not being operated in most cases, in which case they will have been excluded from the study population. This gives us an overestimation of the specificities for both CE and MRI. As with sensitivity, it is not clear which test shows the greater overestimation, as in most studies the authors did not mention which test determined the therapy choice.

Contrary to an earlier report [3], we found a significant influence of the use of FSE MRI techniques, the use of a higher magnetic field, and the use of surface/phased array coil on the results of MRI in both detection of parametrial invasion and advanced disease. This is also what we would expect, as these technical features improve the quality of the images.

Publication bias, which usually appears as an overestimation of test characteristics in smaller studies, was only demonstrated for the evaluation of parametrial invasion using CE as the index test. This may have been caused by the fact that smaller studies were mainly reported by radiologists as the principal investigators. In these studies the use of MRI was studied and explained in detail, whereas the CE results were used in comparisons with the MRI data.

There are several limitations to this meta-analysis. Despite our best efforts to analyse the various factors that could influence our data heterogeneity, we could not exclude the possibility that other factors may have played a role in the diagnostic accuracy of the tests. Unfortunately, the number of studies was too limited to thoroughly examine the effect of co-variables. One could argue not to include old MRI data as these are not comparable with those of recent MRI techniques. However, even with inclusion of studies with old MRI sequences the data are significantly favourable for MRI in comparison to clinical examination. This means that at the present time the sensitivity of MRI is probably higher than found in this meta-analysis. Finally, a wide range of methodological limitations was encountered throughout the various studies, and this shortcoming was more evident in the studies in which CE was assessed.

Our findings suggest that initial CE is insufficient to rule out parametrial invasion or advanced disease and that treatment decisions should be made based on state-of-the-art MRI. An extensive CE should be omitted whenever possible, which has the advantage of avoiding general anaesthesia and potentially reducing health care costs.

On the basis of our systematic review and meta-analysis, and despite the significant heterogeneity among studies, we conclude that MRI is significantly better at ruling out parametrial invasion and advanced disease than CE.

## Appendix. A complete list of the search strategies

### PUBMED

(magnetic resonance imaging[mesh] OR magnetic resonance imag\*[tw] OR nmr [tw] OR mri [tw] OR eua[tw] OR examination under anesth\*[tw] OR examination under anaesth\*[tw] OR examination under general anesth\*[tw] OR examination under local anesth\*[tw] OR clinical examin\*[tw] OR gynecological examin\*[tw] OR gynecological examin\*[tw]) AND (cervic\*[tw] OR cervix\*[tw]) AND (neoplas\*[tw] OR cancer\*[tw] OR tumor[tw] OR tumors[tw] OR tumour\*[tw] OR carcinom\*[tw] OR malign\*[tw]) AND (staging\*[tw] OR stage\*[tw] OR tnm[tw] tumor node metasta\*[tw] OR tumour node metasta\*[tw] OR figo[tw]) AND eng[la] NOT (neck[tw] OR head[tw] OR heada\*[tw] OR oral[tw] OR spine\*[tw] OR spinal[tw] OR intervertebr\*[tw] OR vertebr\*[tw])

### EMBASE

('nuclear magnetic resonance imaging'/exp OR ('magnetic resonance imaging' OR nmr OR mri OR eua OR (('examination under') NEAR/2 (anesth\* OR anaesth\*)) OR ((clinic\* OR gynecologic\* OR gynecologic\*) NEAR/2 examin\*)):ti,ab,de) AND ((cervic\*OR cervix\*) AND (neoplas\* OR cancer\* OR tumor\* OR carcinom\* OR malign\*) AND (staging\*OR stage\* OR tnm OR 'tumor node metastasis' OR 'tumour node metastasis' OR figo)): ti,ab,de AND [english]/lim NOT (neck OR head\* OR oral OR spine\* OR spinal OR intervertebr\* OR vertebr\*):ti,ab,de

## References

1. Pecorelli S, Benedet JL, Creasman WT, Shepherd JH (1999) FIGO staging of gynecologic cancer. 1994-1997 FIGO Committee on Gynecologic Oncology. International Federation of Gynecology and Obstetrics. *Int J Gynaecol Obstet*, 65:243-249.
2. Pecorelli S, Zigliani L, Odicino F (2009) Revised FIGO staging for carcinoma of the cervix. *Int J Gynaecol Obstet*, 105:107-108.
3. Bipat S, Glas AS, van der Velden J, Zwinderman AH, Bossuyt PM, Stoker J (2003) Computed tomography and magnetic resonance imaging in staging of uterine cervical carcinoma: a systematic review. *Gynecol Oncol*, 91:59-66.
4. Hori M, Kim T, Onishi H, et al. (2011) Uterine tumors: comparison of 3D versus 2D T2-weighted turbo spin-echo MR imaging at 3.0 T--initial experience. *Radiology*, 258:154-163.
5. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J (2003) The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol*, 3:25.
6. Arends LR, Hamza TH, van Houwelingen JC, Heijenbrok-Kal MH, Hunink MG, Stijnen T (2008) Bivariate random effects meta-analysis of ROC curves. *Med Decis Making*, 28:621-638.
7. Higgins JP, Thompson SG, Deeks JJ, Altman DG (2003) Measuring inconsistency in meta-analyses. *BMJ*, 327:557-560.
8. Thompson SG, Sharp SJ (1999) Explaining heterogeneity in meta-analysis: a comparison of methods. *Stat Med*, 18:2693-2708.
9. Deeks JJ, Macaskill P, Irwig L (2005) The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol*, 58:882-893.
10. Egger M, Smith GD (1998) Bias in location and selection of studies. *BMJ*, 316:61-66.
11. Macaskill P, Walter SD, Irwig L (2001) A comparison of methods to detect publication bias in meta-analysis. *Stat Med*, 20:641-654.
12. Lien HH, Blomlie V, Iversen T, Trope C, Sundfor K, Abeler VM (1993) Clinical stage I carcinoma of the cervix. Value of MR imaging in determining invasion into the parametrium. *Acta Radiol*, 34:130-132.
13. Qin Y, Peng Z, Lou J, Liu H, Deng F, Zheng Y (2009) Discrepancies between clinical staging and pathological findings of operable cervical carcinoma with stage IB-IIb: a retrospective analysis of 818 patients. *Aust N Z J Obstet Gynaecol*, 49:542-544. \
14. Yang WT, Walkden SB, Ho S, et al. (1996) Transrectal ultrasound in the evaluation of cervical carcinoma and comparison with spiral computed tomography and magnetic resonance imaging. *Br J Radiol*, 69:610-616.
15. Averette HE, Dudan RC, Ford JH, Jr. (1972) Exploratory celiotomy for surgical staging of cervical cancer. *Am J Obstet Gynecol*, 113:1090-1096.
16. Sudarsanam A, Charyulu K, Belinson J, et al. (1978) Influence of exploratory celiotomy on the management of carcinoma of the cervix. A preliminary report. *Cancer*, 41:1049-1053.
17. Maggino T, Bonetto F, Catapano P, Franco F, Valente S, Marchesoni D (1983) Clinical staging versus operative staging in cervical cancer. *Clin Exp Obstet Gynecol*, 10:201-204.
18. Togashi K, Nishimura K, Itoh K (1986) Uterine cervical cancer: Assessment with high-field MR imaging. *Radiology*, 160:431-435.
19. Hricak H, Lacey CG, Sandles LG, Chang YC, Winkler ML, Stern JL (1988) Invasive cervical carcinoma: comparison of MR imaging and surgical findings. *Radiology*, 166:623-631.
20. Waggenspack GA, Amparo EG, Hannigan EV (1988) MR imaging of uterine cervical carcinoma. *J Comput Assist Tomogr*, 12:409-414.
21. Kim SH, Choi BI, Lee HP, et al. (1990) Uterine cervical carcinoma: comparison of CT and MR findings. *Radiology*, 175:45-51.

22. Soeters RP, Beningfield SJ, Dehaeck K, Levin W, Bloch B (1991) The value of magnetic resonance imaging in patients with carcinoma of the cervix (a pilot study). *Eur J Surg Oncol*, 17:119-124.
23. reidler KW, Tamussino K, Szolar DM, Ranner G, Ebner F (1996) Staging of cervical carcinomas. Comparison of body-coil magnetic resonance imaging and endorectal surface coil magnetic resonance imaging with histopathologic correlation. *Invest Radiol*, 31:458-462.
24. Kim MJ, Chung JJ, Lee YH, Lee JT, Yoo HS (1997) Comparison of the use of the transrectal surface coil and the pelvic phased-array coil in MR imaging for preoperative evaluation of uterine cervical carcinoma. *AJR Am J Roentgenol*, 168:1215-1221.
25. Van Vierzen PB, Massuger LF, Ruys SH, Barentsz JO (1998) Fast dynamic contrast enhanced MR imaging of cervical carcinoma. *Clin Radiol*, 53:183-192.
26. Postema S, Pattynama PM, van den Berg-Huysmans A, Peters LW, Kenter G, Trimbos JB (2000) Effect of MRI on therapeutic decisions in invasive cervical carcinoma. Direct comparison with the pelvic examination as a preoperative test. *Gynecol Oncol*, 79:485-489.
27. Hansen MA, Pedersen PH, Andreasson B, Bjerregaard B, Thomsen HS (2000) Staging uterine cervical carcinoma with low-field MR imaging. *Acta Radiol*, 41:647-652.
28. Wang LJ, Wong YC, Chen CJ, Huang KG, Hsueh S (2001) Cervical carcinoma: MR imaging with integrated endorectal/phased-array coils: a pilot study. *Eur Radiol*, 11:1822-1827.
29. Hricak H, Gatsonis C, Chi DS, et al. (2005) Role of imaging in pretreatment evaluation of early invasive cervical cancer: results of the intergroup study American College of Radiology Imaging Network 6651-Gynecologic Oncology Group 183. *J Clin Oncol*, 23:9329-9337.
30. Chung HH, Kang SB, Cho JY, et al. (2007) Can preoperative MRI accurately evaluate nodal and parametrial invasion in early stage cervical cancer? *Jpn J Clin Oncol*, 37:370-375.
31. Javitt MC, Stein HL, Lovecchio JL (1987) MRI in staging of endometrial and cervical carcinoma. *Magn Reson Imaging*, 5:83-92.
32. Togashi K, Nishimura K, Sagoh T, et al. (1989) Carcinoma of the cervix: staging with MR imaging. *Radiology*, 171:245-251.
33. Janus CL, Mendelson DS, Moore S, Gendal ES, Dottino P, Brodman M (1989) Staging of cervical carcinoma: accuracy of magnetic resonance imaging and computed tomography. *Clin Imaging*, 13:114-116.
34. Greco A, Mason P, Leung AW, Dische S, McIndoe GA, Anderson MC (1989) Staging of carcinoma of the uterine cervix: MRI-surgical correlation. *Clin Radiol*, 40:401-405.
35. Sironi S, Belloni C, Taccagni GL, DelMaschio A (1991) Carcinoma of the cervix: value of MR imaging in detecting parametrial involvement. *AJR Am J Roentgenol*, 156:753-756.
36. Kim SH, Choi BI, Han JK, et al. (1993) Preoperative staging of uterine cervical carcinoma: comparison of CT and MRI in 99 patients. *J Comput Assist Tomogr*, 17:633-640.
37. Kaji Y, Sugimura K, Kitao M, Ishida T (1994) Histopathology of uterine cervical carcinoma: diagnostic comparison of endorectal surface coil and standard body coil MRI. *J Comput Assist Tomogr*, 18:785-792.
38. Subak LL, Hricak H, Powell CB, Azizi L, Stern JL (1995) Cervical carcinoma: computed tomography and magnetic resonance imaging for preoperative staging. *Obstet Gynecol*, 86:43-50.
39. Hawighorst H, Knapstein PG, Weikel W, et al. (1996) Cervical carcinoma: comparison of standard and pharmacokinetic MR imaging. *Radiology*, 201:531-539.
40. Scheidler J, Heuck AF, Steinborn M, Kimmig R, Reiser MF (1998) Parametrial invasion in cervical carcinoma: evaluation of detection at MR imaging with fat suppression. *Radiology*, 206:125-129.
41. Hawighorst H, Schoenberg SO, Knapstein PG, et al. (1998) Staging of invasive cervical carcinoma and of pelvic lymph nodes by high resolution MRI with a phased-array coil in comparison with pathological findings. *J Comput Assist Tomogr*, 22:75-81.
42. Yu KK, Hricak H, Subak LL, Zaloudek CJ, Powell CB (1998) Preoperative staging of cervical carcinoma: phased array coil fast spin-echo versus body coil spin-echo T2-weighted MR imaging. *AJR Am J Roentgenol*, 171:707-711.

42. Ng HT, Chen SL, Wang JC, Sheu MH (1998) Preoperative examination with CT, MRI and comparison of both to histopathologic findings in cervical carcinoma. *CME J Gynecol Oncol*, 3:256-257.
43. Shiraiwa M, Joja I, Asakawa T, et al. (1999) Cervical carcinoma: Efficacy of thin-section oblique axial T2-weighted images for evaluating parametrial invasion. *Abdom Imaging*, 24:514-519.
44. Sheu MH, Chang CY, Wang JH, Yen MS (2001) Preoperative staging of cervical carcinoma with MR imaging: a reappraisal of diagnostic accuracy and pitfalls. *Eur Radiol*, 11:1828-1833.
45. Oberoi R, Vohra S, Jain P, Jena A (2002) Staging of carcinoma cervix with MRI and histopathological correlation in 105 cases. *Asian Oceanian J Radiol*, 7:88-94.
46. Choi SH, Kim SH, Choi HJ, Park BK, Lee HJ (2004) Preoperative magnetic resonance imaging staging of uterine cervical carcinoma: results of prospective study. *J Comput Assist Tomogr*, 28:620-627.
47. deSouza NM, Dina R, McIndoe GA, Soutter WP (2006) Cervical cancer: value of an endovaginal coil magnetic resonance imaging technique in detecting small volume disease and assessing parametrial extension. *Gynecol Oncol*, 102:80-85.
48. Fischerova D, Cibula D, Stenhova H, et al. (2008) Transrectal ultrasound and magnetic resonance imaging in staging of early cervical cancer. *Int J Gynecol Cancer*, 18:766-772.
49. Hori M, Kim T, Murakami T, et al. (2009) Uterine cervical carcinoma: preoperative staging with 3.0-T MR imaging--comparison with 1.5-T MR imaging. *Radiology*, 251:96-104.
50. Rotman M, Sedlis A, Piedmonte MR, et al. (2006) A phase III randomized trial of postoperative pelvic irradiation in Stage IB cervical carcinoma with poor prognostic features: follow-up of a gynecologic oncology group study. *Int J Radiat Oncol Biol Phys*, 65:169-176.
51. Sedlis A, Bundy BN, Rotman MZ, Lentz SS, Muddersbach LI, Zaino RJ (1999) A randomized trial of pelvic radiation therapy versus no further therapy in selected patients with stage IB carcinoma of the cervix after radical hysterectomy and pelvic lymphadenectomy: A Gynecologic Oncology Group Study. *Gynecol Oncol*, 73:177-183. d
52. Yeh SA, Wan Leung S, Wang CJ, Chen HC (1999) Postoperative radiotherapy in early stage carcinoma of the uterine cervix: treatment results and prognostic factors. *Gynecol Oncol*, 72:10-15.
53. Waggoner SE (2003) Cervical cancer. *Lancet*, 361:2217-2225.
54. Landoni F, Maneo A, Colombo A, et al. (1997) Randomised study of radical surgery versus radiotherapy for stage Ib-IIa cervical cancer. *Lancet*, 350:535-540.
55. Peters WA, 3rd, Liu PY, Barrett RJ, 2nd, et al. (2000) Concurrent chemotherapy and pelvic radiation therapy compared with pelvic radiation therapy alone as adjuvant therapy after radical surgery in high-risk early-stage cancer of the cervix. *J Clin Oncol*, 18:1606-1613.
56. Rose PG, Bundy BN, Watkins EB, et al. (1999) Concurrent cisplatin-based radiotherapy and chemotherapy for locally advanced cervical cancer. *N Engl J Med*, 340:1144-1153.





# Chapter 3

## **Evaluation of T2-W MR imaging and diffusion-weighted imaging for the early post-treatment local response assessment of patients treated conservatively for cervical cancer: a multicentre study**

Maarten G Thomeer

Vincent Vandecaveye

Loes Braun

Frenchey Mayer

Martine Franckena-Schouten

Peter de Boer

Jaap Stoker

Erik Van Limbergen

Marrije Buist

Ignace Vergote

Myriam Hunink

Helena van Doorn

## Abstract

**Objectives:** To compare MR imaging with or without DWI and clinical response evaluation (CRE) in the local control evaluation of cervical carcinoma after radiotherapy.

**Methods:** In a multicentre university setting, we prospectively included 107 patients with primary cervical cancer treated with radiotherapy. Sensitivity and specificity for CRE and MR imaging (with pre-therapy MR imaging as reference) (2 readers) were evaluated using cautious and strict criteria for identifying residual tumour. Nested logistic regression models were constructed for CRE, subsequently adding MR imaging with and without DWI as independent variables, as well as the pre-treatment post-treatment change in apparent diffusion coefficient (delta ADC).

**Results:** Using cautious criteria, CRE and MR imaging with DWI (reader 1/reader 2) have comparable high specificity (83% and 89%/95%, respectively), whereas MR imaging without DWI showed significantly lower specificity (63%/53%) than CRE. Using strict criteria, CRE and MR imaging with DWI both showed very high specificity (99% and 92%/95%, respectively), whereas MR imaging without DWI showed significantly lower specificity (89%/77%) than CRE. All sensitivities were not significantly different. Addition of MR imaging-DWI to CRE has statistically significant incremental value in identifying residual tumour (reader 1: estimate:1.06, p-value:0.001) (reader 2: estimate:0.62, p-value:0.02) Adding the delta ADC did not have significant incremental value in detecting residual tumour

**Conclusions:** DWI significantly increases the specificity of MR imaging in the detection of local residual tumour. Furthermore, MR imaging with DWI has significant incremental diagnostic value over CRE, whereas adding the delta ADC has no incremental diagnostic value.

### Key points:

1. If MR imaging is used for response evaluation, DWI should be incorporated
2. MR imaging with DWI has diagnostic value comparable/complementary to clinical response evaluation
3. Interreader agreement is moderate to fair for two experienced radiologist readers
4. Quantitative measurements of ADC early post-therapy have limited diagnostic value.

### Abbreviations:

clinical response examination (CRE)

magnetic resonance (MR)

diffusion-weighted imaging (DWI)

T2-weighted MR images (T2-WI)

apparent diffusion coefficient (ADC),

Change in apparent diffusion coefficient ( $\delta$ ADC)

### Key words:

Uterine cervical neoplasm Magnetic resonance imaging Diffusion weighted magnetic resonance imaging, Radiation, Comparative study

## Introduction

Cervical cancer represents a major health burden, with for instance around 12,000 newly diagnosed patients in the United States each year (US Cancer Statistics Working Group 2016). The standard treatment for locally advanced cervical cancer is external beam radiotherapy with image-guided adaptive brachytherapy [1;2] and concurrent chemotherapy [3;4] or hyperthermia [5]. More than 90% of patients show complete local response after such treatment [6-8].

To detect local residual tumour at an early time-point (2-3 months after treatment) and to potentially curative salvage therapy, response assessment after radiotherapy is uniformly recommended [9-11]. Although the effect of salvage surgery on survival is not well studied, retrospective studies suggest that early detection of local residual disease in asymptomatic women may offer survival benefit [12;13]. Moreover, salvage hysterectomy for local disease at a later, i.e., symptomatic stage, has high complication rates or is not feasible due to irresectable disease [14]. Therefore, a well-founded diagnosis of local residual tumour shortly after completing radiotherapy would be helpful in the timely selection of women for salvage surgery, and might contribute to enhanced survival. Secondly, the patient with residual disease not suitable for salvage surgery might be timely informed about her unfortunate prognosis in a timely manner.

In most centres, response assessment is performed via clinical response examination (CRE) at 2-3 months after radiotherapy [1;2]. This pelvic exam can be unsatisfactory due to the altered anatomy after radiotherapy (no cervix, stricture of the vagina, fibrosis) and the discomfort and pain experienced by patients. Subsequently an examination under anesthesia may be necessary.

A seemingly attractive alternative after completing radiotherapy therefore is imaging, either with PET CT scanning or magnetic resonance (MR) imaging.

The use of PET-CT has recently been advised for response evaluation according to the National Comprehensive cancer network (NCCN 2017). In the earlier European Society of Medical Oncology guideline PET-CT is not incorporated for response evaluation [15]. As far as we know, no data exists on its diagnostic value in the first response evaluation. For this reasons we do not use PET-CT in our departments. Further prospective studies are needed to ensure its diagnostic value.

MR imaging has so far been unable to fulfill the need to detect residual disease after radiation due to a false positive rate of up to 45% [16;17] that is mainly attributable to difficulties in differentiating residual tumour from local edema. The development of MR imaging with diffusion-weighted imaging (DWI) has led to investigations of its potential in differentiating

high cellular matrix (as seen in tumour) from low cellular matrix (as seen in edema). Several studies have shown promising results in identifying local residual tumour after radiotherapy, for instance in rectal cancer and breast cancer [18-21]. In 2015, a systematic review described the additional value of DWI in monitoring treatment response in patients undergoing chemoradiotherapy for locally advanced uterine cervical cancer [21]. A greater increase of the apparent diffusion coefficient (ADC) was seen after treatment versus before treatment if response was complete. However, data were not externally validated. Moreover, reference standard was mostly based on MR imaging findings. To the best of our knowledge, only two recent studies have evaluated the accuracy of MR imaging with DWI [22;23]. The first study was a single centre study of retrospective design, while the second study involved retrospective analysis without standard use of pathologic evidence as a reference standard, instead also using radiological follow-up.

The primary aim of our study was to perform a multicentre study of the incremental value of MR imaging with DWI compared to standard MR imaging in the first response evaluation after radiotherapy for cervical cancer. A secondary aim was to evaluate the additional value of MR imaging with or without DWI in comparison to CRE. Histopathology or a minimal follow-up of one year were used as reference standard.

## **Materials and methods**

### **Study design and patients**

This prospective study was approved by the institutional review board in the three participating university hospitals (Erasmus University Medical Centre Rotterdam, Academic Medical Centre Amsterdam, University Hospitals Leuven) and written informed consent was obtained from all patients. The study was registered under trial ID (NTR3345). Study reporting was in accordance with the Standards for Reporting Diagnostic Accuracy Studies (STARD 2015).

From March 2011 to April 2015, all patients of eighteen years and older with a primary tumour of the cervix and scheduled for radiation therapy with curative intent were included. Pregnant women, women with a contraindication for general anesthesia or MRI and women with cognitive deficits were excluded.

### **Clinical response examination**

CRE was based on a pelvic exam, performed by a staff or fellow gynecological oncologist. CRE included visualization of the cervix and vagina with a speculum and bimanual rectovaginal palpation.

Treatment response was scored prospectively as either suspicious or not suspicious for residual tumour and were labeled 'indeterminate' when patients were too difficult to evaluate,

when findings were equivocal, or when the patient refused evaluation at the outpatient clinic. The reader extracting the indeterminate results retrospectively from the medical records was blinded to the MR imaging findings.

For CRE, sensitivity and specificity were calculated based on cautious criteria (i.e., indeterminate results were classified as residual tumour) and on strict criteria (i.e., indeterminate results were classified as complete response).

Due to ethics requirements at one center, the T2-W MRI alone was available to the treating clinicians after CRE. In no patient was the information from DWI used. Where CRE was negative and MR imaging findings without DWI were suspicious for residual tumour, a subsequent clinical examination under anaesthesia was performed or a closer follow-up proposed. Although this could induce verification bias in some cases, it probably had no influence on the reference standard result, since residual tumour would only be detected earlier than during follow-up. In no patient was the information of DWI used.

## MR imaging

In each participating centre, at a predefined time point around 2 months after treatment, imaging was performed on a state-of-the-art 3.0 Tesla MR scanner. In each participating centre, pre-treatment MR imaging was done within 4 weeks of treatment start and response MR imaging was planned 6 to 8 weeks after finalizing radiotherapy. The protocol differed slightly between different vendors (Philips, Best, The Netherlands and General Electrics, Milwaukee, WI). Protocols can be found in Table 1. Imaging was performed before start of treatment and after completing radiotherapy. One institution used intravaginal gel.

## MR Image Analysis

MR imaging was reviewed independently by two experienced readers (reader 1: abdominal radiologist for more than 10 years, reader 2: abdominal radiologist for more than 14 years). Both readers were blinded to clinical and pathological results. The pre-therapy images were at the readers' disposal to correlate the location of the primary tumour, similar to the evaluation process performed in daily clinical practice.

Evaluation was, based on first (1) T2-weighted MR images (T2-WI) in 3 directions, than (2) T2-WI with DWI, and finally (3) the change in Apparent Diffusion Coefficient ( $\delta ADC$ ) between post-therapy and pre-therapy.

The post therapy DWI was analysed first together with the information of before treatment. This is very helpful for localization. The rationale for this was to imitate clinical routine as much as possible. We assume that in clinical practice all patients have an initial scan before therapy.

**Table 1.** The protocol differed slightly between the different vendors (Philips Healthcare, Best, The Netherlands and General Electrics, Milwaukee, Wis). Imaging was performed before the start of treatment and after completing radiotherapy. One institution used intravaginal gel (Figure 2b).

Scanner (3 Tesla)		GE scanner (centre 1)	Philips scanner (centres 2 and 3)
Patients included		85	22
T2-weighting (2D)	TR/TE (ms/ms)	9436/81	4571/80
	Scan plane	3 directions	3 directions
	Flip angle (grades)	120	90
	Matrix size	320 × 320	344× 285
	Bandwidth (kHz)	125	125
	Field view	24	24
	Number of excitations	2	1
	Slice thickness (mm)	4	4
	Gap	0.4	0.4
DWI (2D)	Gradient	3 directions	3 directions
	Scan plane	Axial	Axial
	<i>b</i> values	0–50–100–500–750–1000	0–50–100–500–750–1000
	Bandwidth (kHz)	125	125
	Slice thickness (mm)	4	4

The criterion for residual tumour based on T2-WI was a solid residual mass with intermediate signal intensity. The criterion for residual tumour based on DWI was b1000 hyperintensity, not attributable to T2 shine-through, at the location of the primary tumour as correlated to the T2-WI. For the DWI evaluation, qualitative interpretation of the ADC map provided additional information when deemed necessary by the reader. For example, if the lesion had high intensity on DWI with a corresponding low value on the ADC map (based on qualitative interpretation without quantitative measurement), it was interpreted as tumour. In the third evaluation, based on objective quantitative measurement of the ADC values, the criterion for residual tumour was  $\delta ADC \leq 31.0 \cdot 10^{-4}$ . As published earlier, a  $\delta ADC \geq 62.0 \cdot 10^{-4}$  was considered a complete response [21].

The ADC was measured by circular region of interest (ROI) as large as possible within the solid tumour on the ADC maps on the largest axial tumour slice. Special attention was paid to avoid the areas of focal signal intensity changes, susceptibility artifacts and necrosis. If during response evaluation no tumour was visible, the location of the tumour on the pre-therapy scan was used for ROI measurements. While the justification for this approach is debatable, in most previous studies on cervical carcinoma quantification has been performed in the same way, making it at least comparable [21].

The radiological response of the primary tumour was scored using a five-point score (1= definite complete response, 2= probable complete response, 3= indeterminate result, 4= probable residual tumour, 5= definite residual tumour). As for CRE, sensitivity and specificity were calculated based on cautious criteria (i.e., indeterminate results were classified as residual tumour) and on strict criteria (i.e., indeterminate results were classified as complete response).

## Reference standard

All patients underwent a pelvic examination after radiotherapy and subsequently every 3 months, for at least a year. It is assumed that most small, clinically overt, residual tumours would become obvious within this follow-up period, although some recurrences can occur up to 3 years after therapy. Histopathological evaluation served as the reference standard in case of positive results. A recurrence-free follow-up period of at least one year was considered a surrogate reference standard for a complete local response. Routine histopathology in all cases would have added the burden of an additional EUA and was deemed unacceptable.

## Statistical analysis

For interpretation of T2-weighted images with or without DWI, interobserver agreement on the five-point scale was calculated via Cohen's kappa, using a weighted method based on the squared distance from the diagonal. Kappa values of 0–0.20 were interpreted as poor agreement, 0.21–0.40 as fair, 0.41–0.60 as moderate, 0.61–0.80 as good and 0.81–1 as excellent agreement.

A logistic regression model was constructed first with CRE only (model 1), and subsequently by adding T2-WI (model 2), T2-WI with DWI (model 3) and  $\delta$ ADC (model 4) as independent variables for identification of residual tumour (dependent variable) by the two radiology readers. Our sample size calculation was based on optimizing specificity, in order to decrease the false positive rate due to local edema. We chose an alpha value of 0.05 and estimated the prevalence of residual tumour at 10%. To demonstrate a significant difference in the expected specificity of 0.70 compared to the previously found specificity of 0.55, with a power of 0.80, would require 83 patients without residual tumour. With 10% expected to have residual tumour and 90% a complete response, 92 patients were required. Allowing for 10% lost to follow-up would mean that we needed 101 patients.

## Results

### Patient and response characteristics

One hundred and thirty-three patients were eligible; 26 patients were excluded (refusal of further inclusion (n = 17), change of therapy (n = 5), further therapy elsewhere (n = 2), death during therapy (n = 1), claustrophobia in MR-scanner (n = 1), leaving 107 patients for

analyses. Baseline demographics and clinical characteristics are shown in Table 2. The reason for the conservative treatment of the patient with a FIGO Ia was a relative contraindication for operation.

A complete response was found in 95 patients, whereas twelve patients (11%) had local residual tumour on biopsy or resection (examples Figures 2 and 3). The range of the tumour maximal diameter was between 1 mm and 62 mm, with a mean of 24 mm and a standard deviation of 21 mm. The mean time interval between the end of radiotherapy and response evaluation was 51 days (95% CI 48-53) for CRE and 42 days (95% CI 40-44) for MRI. Follow-up duration in patients without recurrence was between 1.1 year and 5.2 years.

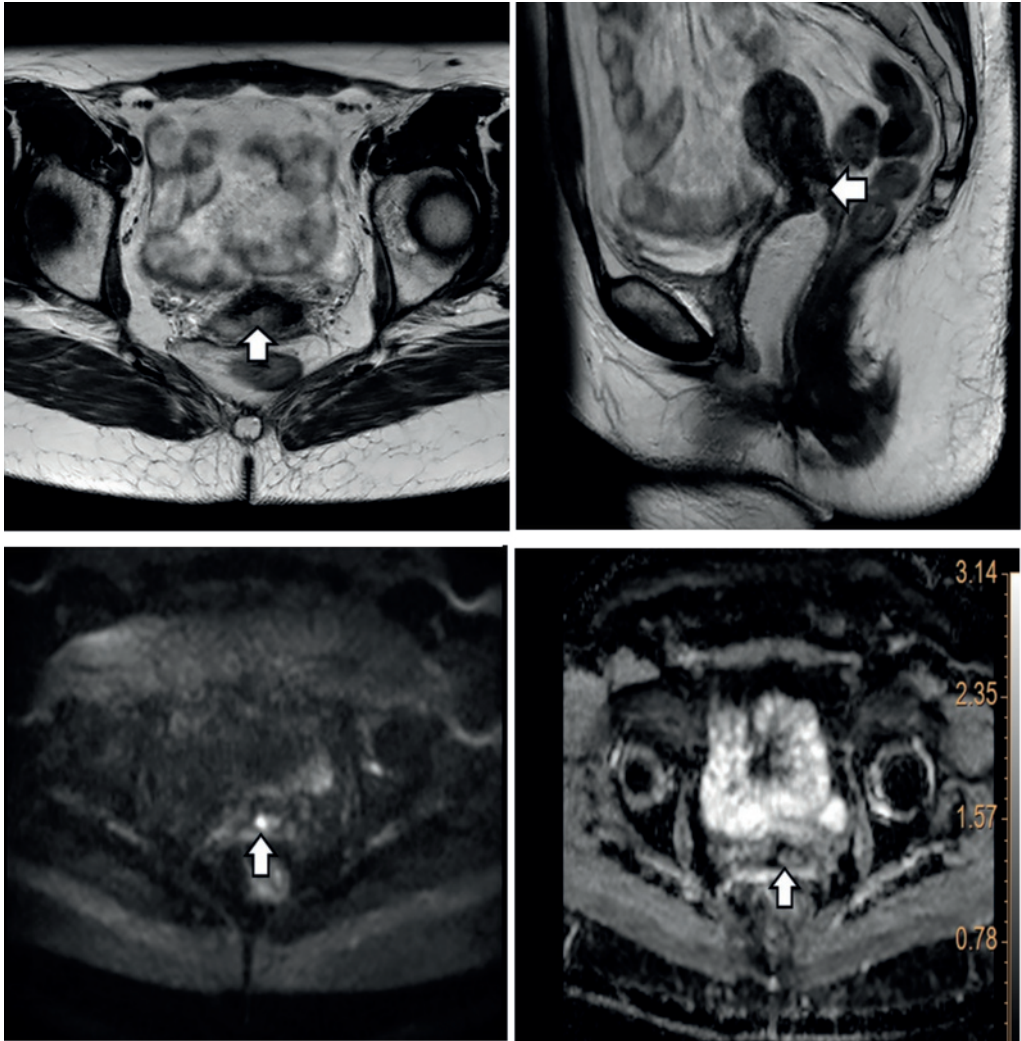


**Figure 1.** This 50-year-old patient was clinically evaluated as non-suspect sixty-one days after radiation therapy. T2-weighted imaging (figure 1a, arrow) showed an indeterminate aspect of the cervix and was therefore supplemented with true cut biopsy under anesthesia. This confirmed residual disease on histopathology. At the end of the study diffusion-weighted imaging (DWI) was analyzed, blinded for the histopathology result. This showed (figure 1b, arrow) a clear hyperintense region. Visual ADC information (low intensity) was used to differentiate from mucosal edema (Figure 1c, arrow), leading to the diagnosis of residual tumour by both readers. The  $\delta ADC$  was only suspect for local residue for reader 2 ( $\delta ADC$ : reader 1:  $52 \times 10^{-4}$  and reader 2:  $29 \times 10^{-4}$ ) based on aforementioned cut-off of  $31 \times 10^{-4}$ .

### *Interobserver agreement*

Interobserver agreement on a five-point scale was fair (0.32) for interpretation of T2-WI and moderate (0.46) for the interpretation of T2-WI + DWI. When the five-point scale was converted to a three-point scale, interobserver agreement was fair for interpretation of T2-WI (0.26) and T2-WI + DWI (0.39).





**Figure 2.** A 30-year-old patient with a normal post-therapeutic clinical examination. T2-weighted MR imaging (Figure 2 a and b, arrow) showed an infracentimetric suspicious region centrally in the cervix. Local follow-up during one year remained normal. Diffusion-weighted imaging (DWI, figure 2c, arrow) was not suspect for both readers. The ADC map (Figure 2d, arrow) was hyperintense (ADC: reader 1:  $132 \times 10^{-4} \text{ mm}^2/\text{s}$  and reader 2:  $125 \times 10^{-4} \text{ mm}^2/\text{s}$ ), confirming the presence of fluid rather than tumour. The delta ADC was not suspect for readers ( $\delta\text{ADC}$ : reader 1:  $63 \times 10^{-4}$  and reader 2:  $52 \times 10^{-4}$ ) based on aforementioned cut-off of  $31 \times 10^{-4}$ . Note the use of intravaginal gel in order to expand the vagina on Figure 2b (white star).

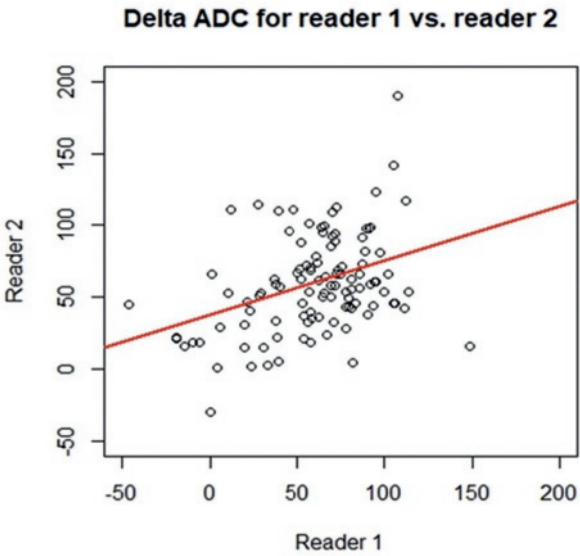
**Table 2.** Patient characteristics.

Patient characteristics		
Patient characteristics		
Total number of patients	107	
Age (median, range)	49 (27–84)	
Tumour stage		
FIGO	Number	%
IA	1	1
IB1	4	4
IB2	13	12
IIA	5	5
IIB	67	63
IIIA	3	3
IIIB	10	9
IVA	4	4
Histologic subtype		
Squamous cell carcinoma	87	81
Adenocarcinoma	13	12
Other	7	7
Treatment		
Chemoradiation	80	75
Radiotherapy + hyperthermia	27	25
(± Induction chemotherapy)		
Brachytherapy	105	98
No brachytherapy	2	2

**Diagnostic performance of CRE and MR imaging**

Diagnostic performance in the interpretation of CRE, T2-WI and T2-WI + DWI are shown in table 3. Almost all specificities of T2-WI with DWI (asterisk) were significantly higher compared to T2-WI. The sensitivities of T2-WI with DWI did not differ significantly compared to T2-WI. Compared to CRE, MR imaging had significantly lower specificities in all situations (bold), and sensitivities and specificities of T2-WI with DWI were not significantly different. Only for reader 1 in the strict setting sensitivity was significantly higher than CRE (bold).

Logistic regression analyses of the effect of adding different MR imaging protocols to the CRE are shown in Table 4. The model with only CRE (model 1) demonstrated the value of CRE ( $\beta=1.74, p=0.0006$ ) in identifying residual tumour. Model 2 demonstrated the significant added value of T2-WI in addition to CRE (reader 1:  $\beta=0.77, p=0.009$ ) (reader 2:  $\beta=0.88, p=0.01$ ). Adding DWI to the T2-WI protocol (model 3) demonstrated the significant added value of DWI (reader 1:  $\beta=1.06, p=0.001$ ) (reader 2:  $\beta=0.62, p=0.02$ ) and showed that T2-WI with DWI can replace CRE. The addition of the  $\delta$ ADC (model 4) had no added value for either reader.



**Figure 3.** Scatter plot comparing delta ADC values for reader1 and reader 2 shows the values were not comparable. ADC: apparent diffusion coefficient

### Difference in ADC between pretreatment and response MRI ( $\delta$ ADC)

We compared  $\delta$ ADC values for patients with a complete response and patients with residual tumour for both readers. According to the scatterplot depicting the  $\delta$ ADC values for both readers (Figure 4), the values of the two readers were not comparable.

**Table 3.** CRE and MR imaging findings of local residual tumour with corresponding diagnostic values using cautious and strict criteria. Using cautious criteria indeterminate findings were labeled as suspicious for residual tumour and using strict criteria indeterminate findings were labeled as not suspicious for residual tumour. CRE: clinical response examination, T2-WI: T2-weighted MR imaging, T2-WI + DWI: T2-weighted imaging + diffusion-weighted imaging. CI: 95% confidence interval. Asterix refers to the outcome which was significantly different on T2-WI + DWI compared to T2-WI. In bold is shown figures which were significantly different on T2-WI or T2-WI + DWI compared to CRE.

		Cautions criteria		Strict criteria	
		Sensitivity	Specificity	Sensitivity	Specificity
CRE		58% (7/12) (CI 32–82)	83% (79/95) (CI 76–89)	25% (3/12) (CI 7–53)	99% (94/95) (CI 95–100)
T2-WI	Reader 1	67% (8/12) (CI 39–88)	<b>63% (60/95)</b> <b>(CI 54–71)</b>	58% (7/12) (CI 32–82)	<b>89% (85/95)</b> <b>(CI 83–94)</b>
	Reader 2	<b>83% (10/12)</b> <b>(CI 56–97)</b>	<b>53% (50/95)</b> <b>(CI 44–61)</b>	75% (9/12) (CI 47–93)	<b>77% (73/95)</b> <b>(CI 69–84)</b>
T2-WI+DWI	Reader 1	<b>83% (10/12)</b> <b>(CI 56–97)</b>	<b>89% (85/95)*</b> (CI 83–94)	<b>83% (10/12)</b> <b>(CI 56–97)</b>	92% (87/95) (CI 85–96)
	Reader 2	50% (6/12) (CI 25–75)	95% (90/95)* (CI 89–98)	50% (6/12) (CI 25–75)	95% (90/95)* (CI 89–98)

\*T2-WI with DWI statistically significantly higher compared to T2-WI

Discussion

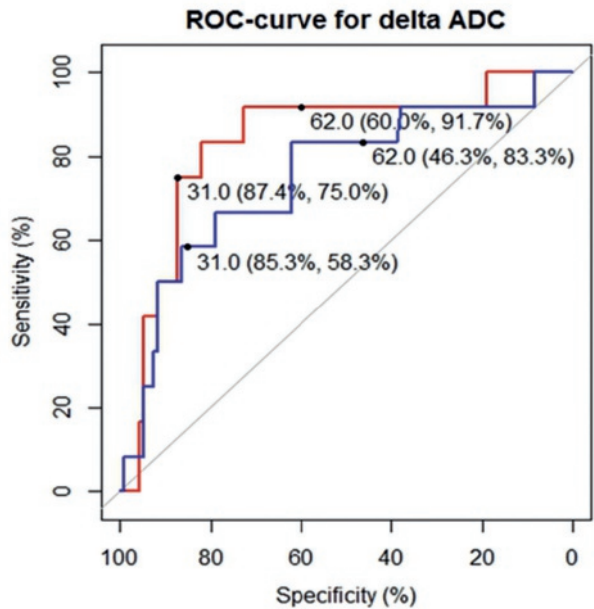
In this prospective multicentre study, we found that DWI significantly increases the specificity of MR imaging in the detection of residual local tumour compared to T2W imaging alone during response evaluation of cervical carcinoma following radiotherapy. MR imaging with DWI also showed significant incremental value over CRE.

In a situation where MR imaging is appended to the treatment response evaluation, DWI information should always be added to reduce the chance of incorrectly declaring an unfavorable outcome, i.e. presumed residual tumour in cases with a complete response. This is in line with our hypothesis, since the presence of edema hampers T2-WI, resulting in misdiagnoses of residual disease. Edema is typically seen in the first months after radiation therapy, at the time when response evaluation is performed to decide on complementary surgery in case of residual disease. Our finding is also in line with a previous study showing that T2-WI alone resulted in up to 50% false positives [16].

**Table 4.** Logistic regression analyses of addition of various MR imaging protocols to the CRE. Parameters with significant p-values are shown in bold. CRE: clinical response examination, T2-WI: T2-weighted MR imaging, T2-WI+ DWI: T2-weighted imaging + diffusion-weighted imaging,  $\delta ADC$  is the difference between the pretreatment ADC and the posttreatment ADC, ADC: apparent diffusion coefficient, estim.: estimate, std err.: standard error. Further ROC analysis demonstrated moderate discrimination (Figure 5): using the aforementioned cut-offs of  $\geq 62 \times 10^{-4}$  for complete response and  $\leq 31 \times 10^{-4}$  for residual tumour, the readers achieved only intermediate sensitivities and specificities.

Reader 1				Reader 2			
<b>Model 1</b>				<b>Model 1</b>			
	Estim.	Std. err.	p-value		Estim.	Std. err.	p-value
Intercept	-2.84	0.46	$p<0.0001$	Intercept	-2.84	0.46	$p<0.0001$
CRE	1.74	0.50	<b>0.0006</b>	CRE	1.74	0.51	<b>0.0006</b>
<b>Model 2</b>				<b>Model 2</b>			
	Estim.	Std. err.	p-value		Estim.	Std. err.	p-value
Intercept	-4.61	0.95	$p<0.0001$	Intercept	-5.52	1.30	$p<0.0001$
CRE	1.22	0.58	<b>0.03</b>	CRE	1.47	0.57	<b>0.01</b>
MRimaging	0.77	0.30	<b>0.009</b>	MRimaging	0.88	0.35	<b>0.01</b>
<b>Model 3</b>				<b>Model 3</b>			
	Estim.	Std. err.	p-value		Estim.	Std. err.	p-value
Intercept	-5.43	1.17	$p<0.0001$	Intercept	-5.36	1.26	$p<0.0001$
CRE	1.08	0.65	0.09	CRE	1.03	0.66	0.1
MRimaging	0.03	0.40	0.9	MRimaging	0.54	0.38	0.2
MRimaging+DWI	1.06	0.32	<b>0.001</b>	MRimaging+DWI	0.62	0.26	<b>0.02</b>
<b>Model 4</b>				<b>Model 4</b>			
	Estim.	Std. err.	p-value		Estim.	Std. err.	p-value
Intercept	-7.09	2.35	0.003	Intercept	-5.08	1.48	0.0006
CRE	1.35	0.77	0.07	CRE	0.96	0.69	0.2
MRimaging	0.10	0.41	0.8	MRimaging	0.56	0.39	0.2
MRimaging+DWI	1.29	0.43	<b>0.003</b>	MRimaging+DWI	0.59	0.29	<b>0.04</b>
$\delta ADC$	0.01	0.02	0.4	$\delta ADC$	-0.00	0.01	0.7

Our findings are in accordance with a recent retrospective study, where ahigh specificity (84%) was found using a combination of high intensity on T2-WI and high intensity on DWI [22]. However, that study differed from ours since their inclusion consisted mainly of lower stages of cervical carcinomas in a neoadjuvant setting.



**Figure 4.** ROC analysis of change of ADC between pretreatment and posttreatment ADC ( $\delta$ ADC) for reader 1 (red line) and reader 2 (blue line). Cut-offs for complete response ( $\geq 62 \times 10^{-4}$ ) and residual tumour ( $\leq 31 \times 10^{-4}$ ) and their corresponding sensitivities and specificities are shown.

To the best of our knowledge, our study is the first to compare CRE with combined T2-WI -DWI for early response assessment after radiation for cervical carcinoma. One study evaluating CRE versus conventional MR imaging in the response assessment after radiation therapy reported that MRI scored significantly worse than CRE, mainly in terms of specificity [16]. However, as that study did not incorporate DWI, it is difficult to compare their results with ours. With the current protocol, we found that MR imaging with DWI significantly improves work-up compared to CRE. It is unlikely that there was bias from the same clinician doing the clinical assessment as doing the CRE, as these assessments were done a year apart.

One could argue that due to current optimized radiation therapy schemes, including MR imaging or CT-guided brachytherapy, the percentage of patients with local residual tumour after radiotherapy treatment for cervical cancer is too low to justify further protocols such as MR imaging. Distant recurrence probably has more impact on survival. However, it remains a fact that only local residual tumour can potentially be cured, for instance with salvage surgery. Despite the low prevalence of local residual tumour without distant metastases, it is therefore imperative that the patient with residual tumour is recognized soon after treatment while salvage is still an option.

Logistic regression analysis showed that if both CRE and MR imaging with DWI are used, only the latter has a significant impact on identifying local residual tumour. A first option would be to replace all CRE by MR imaging with DWI. Although theoretically our data suggest that this would be an acceptable alternative, in clinical practice clinicians would still prefer to examine the patient during the first response evaluation, since it is inexpensive. Moreover, very small superficial and mucosal based residual tumours will probably remain undetected with MR imaging. A second option could be to use MR imaging with DWI in case CRE is difficult to perform or when the patient is uncooperative. A third option is to perform both CRE and MR imaging during the first response evaluation in all cases. The latter two options have four clear advantages. First, the MR images can be reviewed at a later time point and can be discussed with colleagues, which is impossible with CRE. CRE also depends on the experience of the physician. A second advantage is that these initial MRI data are available during follow-up, particularly in cases in which new (clinical) problems and queries arise. The third advantage of using MR imaging is to guide eventual biopsy more precisely in the deeper cervical regions that are more difficult to evaluate with CRE. Finally, in the case of proven local residual tumour, MR imaging is usually better than CRE in identifying local parametrial or pelvic wall invasion, which could make salvage hysterectomy unfeasible, as was shown in an earlier meta-analysis by our study group [24].

A final option is to add MR imaging depending on the probability of local residual tumour and distal tumour recurrence. Higher stages of cervical carcinoma, especially those with possible lymphadenopathy outside the radiation field would have less benefit of salvage surgery, making MRI with DWI less useful.

Quantitative evaluation of the ADC post-therapy versus pre-therapy was less helpful in our study, in contrast to findings reported in an earlier systematic review [21]. The overall absolute cut-offs found for residual tumour versus absence thereof was not reproduced in our multicentre setting. The use of different MR systems may have possibly played a role in our varied results. Indeed, a previous study has shown that absolute ADC values can be vendor-specific, depending on the region scanned in the upper abdomen [25]. As far as we know, the reproducibility of ADC values in the cervical regions across different scanners has not been fully studied. However, in terms of our findings, it is debatable whether the use of different vendors explains the differences, since only two protocols were used, with very similar sequence details and b-values. Based on our findings, we conclude that measurement of ADC differences is time-consuming and is not particularly helpful, even more so because our logistic regression model showed no influence of these data on the final diagnosis of local residual tumour for either radiologist reader.

There are some drawbacks to our study. The interreader agreement was only moderate to fair for MR imaging evaluation, which might suggest that the results are not fully reproducible. This could be an important obstacle to further implementation of this technique. However,



the readers took different approaches to image evaluation despite having received the same predetermined interpretation criteria as mentioned earlier. Reader 1 was more cautious and evaluated indeterminate lesions more often as residual tumour, thereby increasing sensitivity, while reader 2 used more strict criteria for diagnosing residual tumour, i.e., he evaluated indeterminate lesions more often as a complete response. Furthermore, reader 1 had more experience with DWI in the pelvis (10 years versus 6 years), which could also explain his higher sensitivity (although not significantly higher). Further studies should be performed to establish the reproducibility of DWI for identifying residual cervical cancer post radiotherapy.

The results concerning sensitivity were not significantly different across protocols. Due to low prevalence of residual disease, our study would have required over 240 patients to allow conclusions about sensitivity. For reasons of feasibility we powered the study to demonstrate an increase in specificity by adding DWI to the T2-weighted imaging.

According to the most recent guidelines treatment evaluation should be done around 3 months after end of treatment [15]. We decided to perform the MRI some weeks earlier, mainly because we wanted to diagnose residual tumour as soon as possible, in order to enhance the possibilities of local salvage. Too soon after treatment though, the interpretation could be hampered by oedema. However, we prioritized early detection, as the use of DWI for identifying tumour is not compromised by local oedema.

In conclusion, the addition of DWI to the standard MRI protocol significantly improves the value of MR imaging. Furthermore, MR imaging with DWI has significant incremental value relative to CRE in finding residual cervical tumour after radiation therapy, and could possibly replace the latter in situations where CRE is less practical. ADC has no incremental diagnostic value.

## Acknowledgements

We would like to acknowledge the staff of the trial office of the Department of Radiology of Erasmus MC, particularly Laurens Groenendijk and Caroline Van Bavel. Without their daily help in data management this paper would not have been possible. Also, special gratitude to Willemien Hompus (+) for her very helpful assistance during the inclusions.

## Funding

This study has received funding by Fonds Nuts Ohra.



## References

1. Hellebust TP, Kirisits C, Berger D et al (2010) Recommendations from Gynaecological (GYN) GEC-ESTRO Working Group: considerations and pitfalls in commissioning and applicator reconstruction in 3D image-based treatment planning of cervix cancer brachytherapy. *Radiother Oncol* 96:153-160
2. Gaffney DK, Erickson-Wittmann BA, Jhingran A et al (2011) ACR Appropriateness Criteria(R) on Advanced Cervical Cancer Expert Panel on Radiation Oncology-Gynecology. *Int J Radiat Oncol Biol Phys* 81:609-614
3. Vale CL, Tierney JF, Davidson SE, Drinkwater KJ, Symonds P (2010) Substantial improvement in UK cervical cancer survival with chemoradiotherapy: results of a Royal College of Radiologists' audit. *Clin Oncol (R Coll Radiol)* 22:590-601
4. Green J, Kirwan J, Tierney J et al (2005) Concomitant chemotherapy and radiation therapy for cancer of the uterine cervix. *Cochrane Database Syst Rev*. 10.1002/14651858.CD002225.pub2:CD002225
5. Lutgens L, van der Zee J, Pijls-Johannesma M et al (2010) Combined use of hyperthermia and radiation therapy for treating locally advanced cervical carcinoma. *Cochrane Database Syst Rev*. 10.1002/14651858.CD006377.pub2:CD006377
6. Lindegaard JC, Fokdal LU, Nielsen SK, Juul-Christensen J, Tanderup K (2013) MRI-guided adaptive radiotherapy in locally advanced cervical cancer from a Nordic perspective. *Acta Oncol* 52:1510-1519
7. Potter R, Georg P, Dimopoulos JC et al (2011) Clinical outcome of protocol based image (MRI) guided adaptive brachytherapy combined with 3D conformal radiotherapy with or without chemotherapy in patients with locally advanced cervical cancer. *Radiother Oncol* 100:116-123
8. Hellebust TP, Kristensen GB, Olsen DR (2010) Late effects after radiotherapy for locally advanced cervical cancer: comparison of two brachytherapy schedules and effect of dose delivered weekly. *Int J Radiat Oncol Biol Phys* 76:713-718
9. Committee on Practice B-G (2002) ACOG practice bulletin. Diagnosis and treatment of cervical carcinomas, number 35, May 2002. *Obstet Gynecol* 99:855-867
10. Elit L, Kennedy EB, Fyles A, Metser U (2016) Follow-up for cervical cancer: a Program in Evidence-Based Care systematic review and clinical practice guideline update. *Curr Oncol* 23:109-118
11. Elit L, Fyles AW, Devries MC, Oliver TK, Fung-Kee-Fung M, Gynecology Cancer Disease Site G (2009) Follow-up for women after treatment for cervical cancer: a systematic review. *Gynecol Oncol* 114:528-535
12. Bodurka-Bevers D, Morris M, Eifel PJ et al (2000) Posttherapy surveillance of women with cervical cancer: an outcomes analysis. *Gynecol Oncol* 78:187-193
13. Chou HH, Wang CC, Lai CH et al (2001) Isolated paraaortic lymph node recurrence after definitive irradiation for cervical carcinoma. *Int J Radiat Oncol Biol Phys* 51:442-448
14. Sardain H, Lavoue V, Redpath M, Bertheuil N, Foucher F, Leveque J (2015) Curative pelvic exenteration for recurrent cervical carcinoma in the era of concurrent chemotherapy and radiation therapy. A systematic review. *Eur J Surg Oncol* 41:975-985
15. Colombo N, Carinelli S, Colombo A et al (2012) Cervical cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 23 Suppl 7:vii27-32
16. Vincens E, Balleyguier C, Rey A et al (2008) Accuracy of magnetic resonance imaging in predicting residual disease in patients treated for stage IB2/II cervical carcinoma with chemoradiation therapy : correlation of radiologic findings with surgicopathologic results. *Cancer* 113:2158-2165
17. Hequet D, Marchand E, Place V et al (2013) Evaluation and impact of residual disease in locally advanced cervical cancer after concurrent chemoradiation therapy: results of a multicenter study. *Eur J Surg Oncol* 39:1428-1434
18. Kim SH, Lee JM, Hong SH et al (2009) Locally advanced rectal cancer: added value of diffusion-weighted MR imaging in the evaluation of tumor response to neoadjuvant chemo- and radiation therapy. *Radiology* 253:116-125

19. Lambregts DM, Vandecaveye V, Barbaro B et al (2011) Diffusion-weighted MRI for selection of complete responders after chemoradiation for locally advanced rectal cancer: a multicenter study. *Ann Surg Oncol* 18:2224-2231
20. Bui E, Belli P, Costantini M et al (2015) Role of the Apparent Diffusion Coefficient in the Prediction of Response to Neoadjuvant Chemotherapy in Patients With Locally Advanced Breast Cancer. *Clin Breast Cancer* 15:370-380
21. Schreuder SM, Lensing R, Stoker J, Bipat S (2015) Monitoring treatment response in patients undergoing chemoradiotherapy for locally advanced uterine cervical cancer by additional diffusion-weighted imaging: A systematic review. *J Magn Reson Imaging* 42:572-594
22. Jalaguier-Coudray A, Villard-Mahjoub R, Delouche A et al (2017) Value of Dynamic Contrast-enhanced and Diffusion-weighted MR Imaging in the Detection of Pathologic Complete Response in Cervical Cancer after Neoadjuvant Therapy: A Retrospective Observational Study. *Radiology*. 10.1148/radiol.2017161299:161299
23. Park JJ, Kim CK, Park BK (2016) Prediction of disease progression following concurrent chemoradiotherapy for uterine cervical cancer: value of post-treatment diffusion-weighted imaging. *Eur Radiol* 26:3272-3279
24. Thomeer MG, Gerestein C, Spronk S, van Doorn HC, van der Ham E, Hunink MG (2013) Clinical examination versus magnetic resonance imaging in the pretreatment staging of cervical carcinoma: systematic review and meta-analysis. *Eur Radiol* 23:2005-2018
25. Donati OF, Chong D, Nanz D et al (2014) Diffusion-weighted MR imaging of upper abdominal organs: field strength and intervendor variability of apparent diffusion coefficients. *Radiology* 270:454-463

# Chapter 4

Can MR imaging at 3.0 Tesla reliably  
detect patients with endometriosis?  
Initial results

Maarten G Thomeer

Anneke B Steensma

Evert J van Santbrink

Francois E Willemsen

Piotr A Wielopolski

Myriam G Hunink

Sandra Spronk

Joop S Laven

Gabriel P Krestin

## Abstract

**Objective:** To determine whether an optimized 3.0 Tesla MR imaging protocol is sensitive and specific enough to detect patients with endometriosis.

**Material and methods:** Prospective cohort study with consecutive patients. Forty consecutive patients with clinical suspicion of endometriosis underwent 3.0 Tesla MR imaging including a T2-weighted high-resolution fast spin echo sequence (spatial resolution (SR) =  $0.75 \times 1.2 \times 1.5 \text{ mm}^3$ ) and a 3D T1-weighted high-resolution gradient echo sequence (SR =  $0.75 \times 1.2 \times 2.0 \text{ mm}^3$ ). Two radiologists reviewed the data set with consensus reading. During laparoscopy, which was used as reference standard, all lesions were characterized according to the revised criteria of the American Fertility Society. Patient-level and region-level sensitivities and specificities and lesion-level sensitivities were calculated.

**Results:** Patient-level sensitivity was 42% for stage I (5/12) and 100% for stage II, III and IV (25/25). Patient-level specificity for all stages was 100% (3/3). The region-level sensitivity and specificity was 63% and 97% respectively. The sensitivity per lesion was 61% (90% for deep lesions, 48% for superficial lesions and 100% for endometriomata). The detection rate of obliteration of the Cul-the-sac was 100% (10/10) with no false positive findings. The interreader agreement was substantial to perfect (kappa = 1 per patient, 0.65 per lesion and 0.71 for obliteration of the Cul-the-sac).

**Conclusions:** An optimized 3.0 Tesla MR imaging protocol is accurate in detecting stage II to stage IV endometriosis.

### Key words

MRI, Endometriosis, 3.0 Tesla

## Introduction

Endometriosis is characterized by development of ectopic endometrial tissue causing cyclic abdominal pain and subfertility [1]. The incidence of endometriosis in the general population is estimated to be 12% to 45% [2].

Development of a non-invasive test for endometriosis has long been an important priority in endometriosis research because the delay that precedes an accurate diagnosis of endometriosis seems to be substantial [3;4]. Since most of these patients are in their reproductive life span but also may have significant pain symptoms, this will have a significant influence on fertility chances and quality of life. The reference standard for the diagnosis of endometriosis currently still is a diagnostic laparoscopy, an invasive procedure with known complications [5].

So far, several non-invasive tests have been developed for detecting patients with endometriosis. Inflammatory markers in either the peritoneal fluid or blood as evidence of induction of systemic inflammatory responses in women with endometriosis are depicted with limited success [6-9]. Similarly, tests with tumor markers like Ca125 and Ca19-9 showed disappointing results so far [7-11]. Ultrasonography can only detect deep lesions and therefore is unreliable to depict lower stages, which are mainly characterized by superficial lesions [12;13].

Recent literature shows that MR imaging using a 3.0 Tesla (T) scanner can reliably detect deep endometriosis [14]. Studies on MR imaging of endometriosis so far have mainly focused on the preoperative setting in order to map the locations of deep endometriosis. Such mapping allows the gynecologist to plan the operation, should this be performed. However, the use of MR imaging to make the diagnosis of endometriosis irrespective of the presence and extent of deep or superficial endometriosis has not fully been addressed until now.

The purpose of this study is to explore whether an optimized 3.0 Tesla MR imaging protocol for endometriosis may be sufficiently sensitive to detect the disease in a clinical setting, compared to laparoscopic exploration.

## Material and Methods

### Patients

The study protocol was approved by the institutional medical ethical review board (MERB). Sixteen patients gave written informed consent and for the other 24 patients informed consent was waived by the MERB since they underwent MRI for clinical purpose (preoperative imaging). The consecutive group of 40 patients with clinical suspicion of endometriosis were

scheduled to undergo laparoscopy in a period of 2 years from november 2010 to december 2012. Exclusion criteria were the use of contraceptives or hormonal suppressive medication (like gonadotropin-releasing hormones), contraindication to MR imaging (pacemaker, different metallic bodies, claustrophobia), age younger than 18 and postmenopausal status.

Patients were included from the division of Gynecology for pain complaints and from the division of Reproductive Medicine with subfertility as major complaint and with specific symptoms suspicious for endometriosis.

## MR imaging

After inclusion, the MR imaging was performed between the ninth and fifteenth day of the menstrual cycle. Buscopan® (butylscopolamine 40 mg, Boehringer Ingelheim, Ingelheim am Rhein, Germany) was administered intravenously. In order to suppress bowel motion during at least half an hour Buscopan® was administered at double dose (40 mg instead of 20 mg) in a very slow drip-infusion of 50 cc Natrium chloride. No intravenous MR imaging contrast was used.

Scans were acquired on a 3.0 Tesla scanner (GE Healthcare, Milwaukee, USA) using a whole-body radio frequency (RF) coil transmitter and an 8-channel cardiac phased array coil for signal reception. After a 3-plane localizer, a two-dimensional (2D) T2-weighted fast spin echo (FSE) sequence was performed in the axial orientation with the following imaging parameters: TR/TE = 12000/70 ms; echo train length (ETL) = 10; readout bandwidth (BW)= 50 kHz, number of excitations (NEX) = 2, 100 sections; spatial resolution =  $0.75 \times 1.2 \times 1.5 \text{ mm}^3$ ; field of view (FOV) =  $24 \times 24 \text{ cm}^2$ . The acquisition time was around 8 minutes.

The examination was followed by a axial three-dimensional (3D) fat-suppressed T1-weighted fast RF spoiled gradient echo (SPGR) sequence. The imaging parameters used were: TR = 13/1.3 ms; flip angle =  $30^\circ$ ; BW = 31.2 kHz, NEX = 2, 80 sections, spatial resolution =  $0.75 \times 1.2 \times 2.0 \text{ mm}^3$ , FOV =  $24 \times 24 \text{ cm}^2$ . The acquisition time was around 7 minutes.

## Laparoscopy

Within 2 months after the MR imaging scan was performed, all patients underwent a diagnostic laparoscopy. Laparoscopic findings were used as reference standard to evaluate the MR findings. The location and size of each peritoneal lesion and endometrioma found during laparoscopy was recorded with digital video (avi) and reviewed by two experienced gynecologists (BLINDED). Both gynecologists were highly qualified and had extensive practical experience with laparoscopy and especially qualified in detecting endometriosis. Inter-observer agreement with consensus reading was performed. The pelvic locations include the cul-the-sac, uterus, torus uterinus, recto-sigmoid, bladder, pelvic side walls, right uterosacral ligament, left uterosacral ligament, left ovary, right ovary, and their respective ovarian fossa.

Lesions were categorized by depth according to the revised American Fertility Society (rAFS) criteria. For the approximation of the depth of the lesion categorization according to Stratton et al. was used with different depths (< 2 mm, 2–4 mm, > 4mm) [15]. A lesion was classified as superficial if it was stated as no deeper than 4 mm below the peritoneal surface [15;16]. Because this is difficult to assess a consensus reading was performed. On a clarifying drawing and in a table the gynecologists, blinded for the MR imaging findings, were asked to write the location of the lesions.

Obliteration of the Cul-the-sac was mentioned if the Cul-the-sac was not accessible from the peritoneal space, due to the presence of adhesions.

Finally, patients were categorized according to the revised criteria of the American Fertility Society (rAFS) in stage I(minimal), II(mild), III (moderate) and IV(severe).

### Image analysis

Two experienced radiologists (BLINDED), with 13 and 12 years of experience in abdominal MR imaging, analyzed independently and blindly the data from the MR imaging on a PACS workstation (iSITE, Philips, Eindhoven, The Netherlands). They had no information regarding clinical data. On a clarifying drawing and in a table they were asked to identify the location of the lesions.

The diagnosis of an endometrioma in the ovary was based on shading on T2-weighted images and hyperintensity on T1-weighted images [17]. Fibrotic-like tissue on T2-weighted images was stated as deep endometriosis. Focal T1-weighted hyperintense foci without T2-weighted abnormalities were considered as superficial endometriotic lesions [17].

The presence of obliteration without a definable mass is based on the visibility of adhesions between the uterus and bowel loops as previously mentioned by Kataoka et al [18].

### Statistical analysis

The diagnostic performance of MR imaging for the detection of patients with endometriosis was compared with the findings on laparoscopy, the reference standard test. The estimated measures were sensitivity, specificity, positive predictive value and negative predictive value , and were calculated per patient and per region separately. These diagnostic parameters were expressed with a 95% confidence interval (CI) for per-region analysis and 97.5% CI using exact test for the per-patient analysis. The per-patient calculations included classification according to the rAFS stage.

The pelvis was divided into 12 regions similar to those used in the laparoscopy (see above). For the per-region calculations the estimated measures and corresponding 95% confidence intervals (CIs) were calculated by using variance estimates that take into account the correlated

data owing to the multiple regions within a patient and were based on the generalized estimating equations (GEE) method with an independent working correlation matrix [19].

In addition, the sensitivity and positive predictive value were calculated per-lesion (superficial lesions, deep lesions, endometriomata) and for the presence of obliteration of the Cul-the-sac. These diagnostic parameters were also expressed with a 95% CI for per-lesion analysis and 97.5% CI using exact test for obliteration of the Cul-the-sac. Specificity and the negative predictive value are dependent on the number of true negatives, and therefore these parameters can only be meaningfully determined in the per-patient and per-region analysis but not in the per-lesion analysis.

Based on the presence of obliteration, endometriomata, deep lesions and superficial lesions, patients are categorized as having more pronounced endometriosis (higher stages, III and IV) or less pronounced endometriosis (lower stages, I and II). A classification tree was developed to determine the optimal algorithm for classification into the higher and lower stages as compared to the reference standard test in our dataset.

Interobserver agreement was assessed using the kappa value ( $k$ ). A  $k \leq 0.20$  was interpreted as slight agreement; 0.21–0.40, fair agreement; 0.41–0.60, moderate agreement; 0.61–0.80, substantial agreement; and  $\geq 0.81$ , almost perfect agreement [20]. Final results were based on consensus agreement between the two readers.

All statistical analyses were performed by using SPSS 16.0 (SPSS, Chicago, ILL), and STATA 11.0 (STATA, College Station, Tex).

## Results

A total of 40 patients were included in this study with a median age of 25 (range: 18–39) years. During laparoscopy 3 patients were diagnosed without having endometriosis, 12 patients had stage I endometriosis, 8 patients had stage II endometriosis, 2 patients had stage III endometriosis, 15 had stage IV endometriosis. Eighteen endometriomata, 29 deep lesions and 126 superficial lesions were found.

The per-patient calculations of detection of endometriosis by MR imaging compared to laparoscopy as reference standard had a sensitivity of 81% (30/37; 97.5% CI: 65–92%), a specificity of 100% (3/3; 97.5% CI: 29–100%), a positive predictive value of 100% (30/30; 97.5% CI: 88–100%) and a negative predictive value of 70% (7/10; 97.5% CI: 29–100%) for all stages combined. Table 1 shows the sensitivity of detecting endometriosis by stage. There were seven patients with false negative findings on MR imaging, all diagnosed with stage I endometriosis according to the laparoscopic findings. In these seven patients only some



small spots of superficial endometriosis were seen by laparoscopy. In eleven patients, MR imaging found only superficial lesions, which was confirmed by laparoscopy. Four of them were laparoscopically classified as stage I and seven as stage II, based on a combination of superficial lesions, and fibrotic strands.

**Table 1.** Patient-level analysis: Detection of patients with endometriosis according to the stage of endometriosis based on laparoscopy.

	Laparoscopy <sup>+1</sup>				Laparoscopy <sup>-2</sup>
	Stage I	Stage II	Stage III	Stage IV	
MRI <sup>-3</sup>	5	8	2	15	0
MRI <sup>-4</sup>	7	0	0	0	3
Sensitivity	42%	100%	100%	100%	

<sup>1</sup>Patient positive for endometriosis based on consensus reading of two laparoscopists; <sup>2</sup>Patient negative for endometriosis based on consensus reading of two laparoscopists; <sup>3</sup>Patient positive for endometriosis based on consensus reading of two radiologists; <sup>4</sup>Patient negative for endometriosis based on consensus reading of two radiologists.

The per-region calculations of detection of endometriosis by MR imaging using the GEE method had similar estimates as the per-region calculations without GEE, however the 95% CIs were slightly wider: a sensitivity of 63% (95% CI: 54–70%), a specificity of 97% (95% CI: 94–98%) a positive predictive value of 92% (95% CI: 86–96%) and a negative predictive value of 82% (95% CI: 78–85%) (Table 2).

The sensitivity for deep lesions was 90% (26/29; 95% CI: 78–100%) was significantly higher in comparison to smaller lesions (Figure 1) (48%; 61/126; 95% CI: 41–59%) (*P* value < 0.001) (Table 3). Three deep lesions were missed: one was visible retrospectively (sacrouterine ligament), one lesion was in the anterior abdominal wall and outside the scan, and the last one was not clearly visible retrospectively. The positive predictive value for detecting deep lesions was 100% (26/26; 97.5% CI: 87–100) and higher than for superficial lesions (Figure 2) (87 %; 61/70; 95% CI: 79–95%), although not statistically significant (*P* value = 0.06). There were 9 false positive findings in 114 detected lesions (deep, superficial lesions and endometriomata) (Table 3). These false positive findings consisted of one lesion on the bladder dome, three lesions on the sacrouterine ligament, three lesions on the ovariae and two lesions in the fossa ovarica.

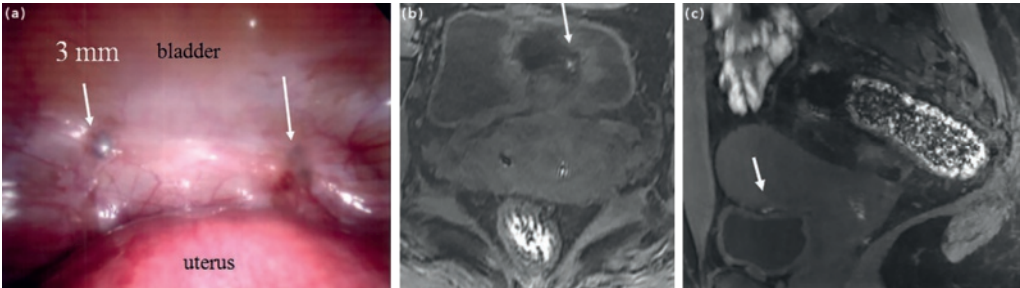
The sensitivity and positive predictive value in detecting lesions unrelated to the size was 61% (105/173; 95% CI: 53–68%) and 92% (105/114; 95 % CI: 87–97%). respectively.

The sensitivity, specificity, positive predictive value and negative predictive value of detecting obliteration of the Cul-the-sac were 100% (10/10; 97.5% CI: 69–100%), 100 % (30/30; 97.5% CI: 88–100%), 100% (10/10; 97.5% CI: 69–100%) and 100% (30/30; 97.5% CI: 88–100%).

**Table 2.** Region-level analysis: Sensitivity, specificity, NPV and PPV according to the 12 regions in the pelvis.

	Bladder	Fossa ov left	Fossa ov right	SUL left	SUL right	Torus	Rectum	Douglas	Pelvic wall	Ov left	Ov right	Uterus
Sensitivity	0.65	0.65	0.59	0.56	0.53	0.86	0.88	0.5	0.87	1	1	0.9
Specificity	1	0.92	0.92	0.92	0.92	0.92	1	1	1	1	1	0.96
NPV	0.53	0.65	0.61	0.6	0.57	0.96	0.96	0.65	0.92	1	1	0.96
PPV	1	0.92	0.91	0.91	0.91	0.75	1	1	1	1	1	0.9

NPV = negative predictive value; Ov = ovarian; PPV = positive predictive value; SUL = sacrouterine ligament.

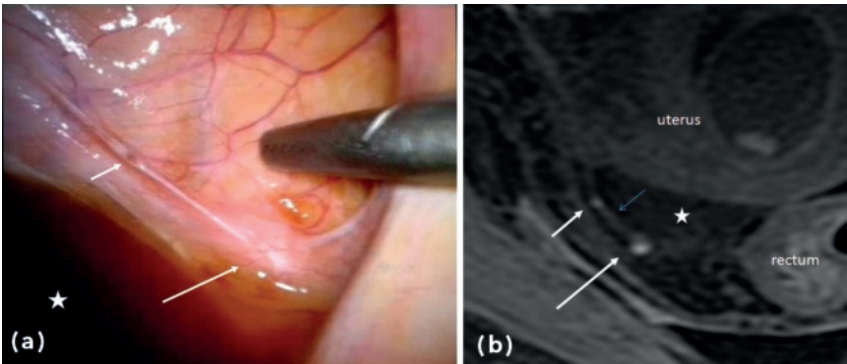


**Figure 1.** Laparoscopic view and MR images of a patient with endometriosis in the plica vesico-uterina. **(a)** endoscopic view with two brown superficial lesions on the bladder with both an estimated depth less than 4 mm. **(b)** Axial high resolution 3D T1-weighted gradient image of the same patient showing an hyperintense spot on the bladder dome (eg. plica vesico-uterina). Due to partial volume effect this lesion is mainly surrounded by peritoneal fat. **(c)** Reconstructed sagittal high resolution 3D T1-weighted gradient image of the same patient visualizes the second lesion also in the plica vesico-uterina.

**Table 3.** Lesion-level analysis: Detection of lesions according to laparoscopic findings based on lesion depth and presence of ovarian lesions (endometriomata)

	Laparoscopy <sup>+1</sup>				Laparoscopy <sup>-2</sup>
	< 2 mm	2–4 mm	> 4 mm	Ovarian lesion	
MRI <sup>+3</sup>	24	37	26	18	9
MRI <sup>-1</sup>	33	32	3	0	n/a
Sensitivity	42%	54%	90%	100%	

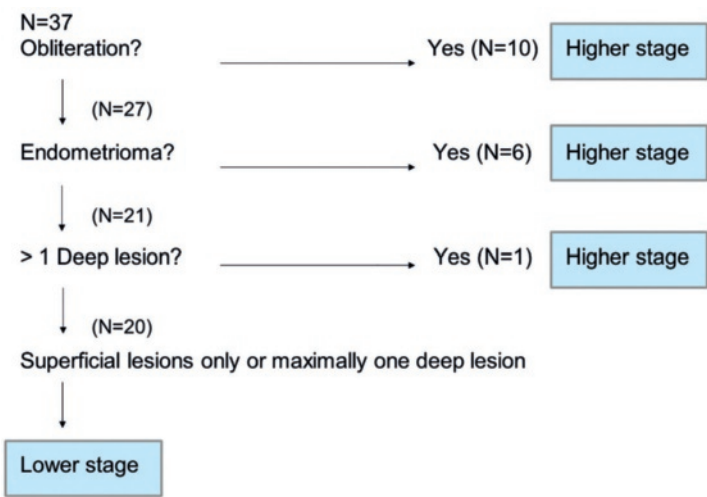
<sup>1</sup>Per lesion positive findings based on consensus reading of two laparoscopists. <sup>2</sup>Per lesion negative findings based on consensus reading of two laparoscopists. <sup>3</sup>Per lesion positive for finding based on consensus reading of two radiologists.



**Figure 2.** Laparoscopic view and MR images of a patient with endometriosis in de Cul-the-sac. **(a)** endoscopic view shows two spots of superficial endometriosis in the pelvic sidewall with a fibrotic strand. The larger lesion is white (right arrow) and the very small lesion is brown (left arrow). **(b)** Axial high resolution 3D T1-weighted gradient image of the same region shows very superficial bright spots in de Cul-the sac (large white arrow). Asterisk = ascites; small blue arrow = peritoneal surface.

The classification tree of patients having lower stage endometriosis (I and II) versus higher stage endometriosis (III and IV) is reflected in an algorithm (Figure 3). In the presence of obliteration, 1 endometrioma or more than 1 deep lesion on MR imaging, a patient is classified as having higher stage of endometriosis.

The per-patient inter-observer agreement for radiologists was perfect ( $k = 1$ ). The per-lesion inter-observer agreement was substantial ( $k = 0.65$ ). The inter-observer agreement for the detection of obliteration was substantial ( $k = 0.71$ ). The per-region inter-observer was not measured since the data for the per-region analysis were derived from the consensus findings per lesion.



**Figure 3.** Classification tree. Explorative algorithm in order to classify patients as having lower stage (stage I or II) or higher stage (stage III or IV) endometriosis on MR imaging. (> 1 deep lesion: more than one deep lesion)

## Discussion

The current study provides evidence that MR imaging is accurate in detecting patients with mild, moderate or severe endometriosis (stage II-IV) with perfect interreader agreement. In patients with minimal endometriosis (stage I) MR imaging is unable to detect the disease.

During laparoscopy endometriosis is classified according the rAFS into four different stages: minimal, mild, moderate and severe (stage I to IV) [16]. It is of importance to detect all stages of endometriosis, including the lower stages, as pain and infertility are not always related to the severity of the disease as classified by the rAFS. Regarding endometriosis related pain

conditions, it is known that patients can suffer from severe dysmenorrhea caused by mild endometriosis and that complaints can be minimal in cases of severe endometriosis [21]. Concerning subfertility, it is known that all stages can cause fecundity problems. Although higher stages probably cause more infertility, detection of the lower stages are very important since fertility may be improved with surgical intervention mainly in this subgroup, as showed by a recent meta-analysis [22].

Stage III and IV endometriosis according to the AFS classification reflects usually deep endometriosis, endometriomata or Cul-the-sac obliteration. Our results in detecting deep endometriosis and endometriomata are in concordance with previously reported studies [14; 17]. The largest study is from Bazot et al, where they detected around 90% deep lesions using a 1.5 Tesla scanner [17].

As a consequence of our findings, the detection of higher stages of endometriosis (stage III and IV) can be performed with high sensitivity since these are mainly affected by deep lesions, endometriomata, or Cul-the-sac obliteration. Our protocol utilizing a 3.0 Tesla MR scanner seems to be an improvement in comparison to a previous protocol using 1.5 Tesla MR scanner which had lower resolution [15]. In the previous study using a 1.5 Tesla scanner, the investigators detected only 10 of 15 patients with stage III and IV endometriosis, whereas we had no false negatives.

It is also of importance to depict the lower stages I and II of endometriosis. These lower stages reflect less deep lesions and more superficial endometriosis. On MR imaging superficial lesions are mainly detected by bright spots on T1-weighted sequences, corresponding to local blood residues on the peritoneal surface. Earlier reports found a very low sensitivity in depicting superficial lesions (between 5 and 15%) [15;17]. Our results were better, but still only 50% of the superficial lesions were detected. Even though we detected only half of all superficial lesions, we were able to make the diagnosis in all patients with stage II endometriosis, including 7 patients with only superficial endometriosis. The detection of minimal endometriosis, reflecting sporadic superficial lesions (stage I), is still unsatisfactory.

The lower sensitivity of previous studies, regarding superficial endometriosis, might be related, apart from the lower scan field, to the use of a 2D technique for T1-weighted imaging. With a 2D-technique, small vessels in the pelvis can be bright due to slow flow phenomenon, hence making it very difficult to differentiate it from superficial endometriosis spots. Using a 3D gradient sequence, these pseudo-lesions are optimally suppressed. This provided us the opportunity to classify every bright spot on T1-weighted images as endometriosis with only 9 false positive findings in a total of 174 detected lesions and a specificity of 97% based on a per region analysis. Note that specificity has no meaning in the lesion-level analysis.

To be able to predict whether a patient has a lower stage of endometriosis (stage I and II) versus a higher stage of endometriosis (stage III and IV) an explorative algorithm was developed. The cut off between lower versus higher stage of endometriosis was arbitrarily chosen because of the small data set. Not surprisingly, Cul-the-sac obliteration on MR imaging strongly predicted the presence of higher stage endometriosis. Furthermore, the presence of an endometrioma or at least 2 deep lesions was found indicative for a higher stage of endometriosis. We emphasize that this algorithm could be useful to stage patients based on MR imaging but it needs to be validated in a larger patient population.

Our findings, although preliminary, could have impact for clinical practice, in patients with subfertility and pain. In the case of subfertility it would be very useful to have a non-invasive test which could detect endometriosis. Instead of a diagnostic laparoscopy in order to rule out endometriosis, this could result in an advice to add MR imaging as a tool to detect endometriosis early in the fertility screening on a routine basis.

Another implication for MR imaging is its use in pain related conditions. Endometriosis can cause serious morbidity. Complaints can often be atypical with a doctor delay as consequence. A survey from 2006 showed that the mean diagnostic delay of patients suffering from dysmenorrhea is approximately 11 years, and that most patients had visited more than 5 doctors before a final diagnosis could be made [23]. A non-invasive test like MR imaging could lead to treatment in an earlier stage of the disease and hence prevent or slow-down progression of the disease and symptoms.

A limitation of our study is the relatively small patient group and the patient selection based on clinical suspicion of having endometriosis. Our study indicates that it is possible to detect all stages of endometriosis except minimal endometriosis. Using this protocol it has to be analyzed if the results can be replicated in a larger patient population and in patients with a less strong clinical suspicion of having endometriosis.

Our patient group was also limited to patients without hormonal suppression. This was performed in order to enhance the visibility of active lesions on T1-weighted images. In the fertility screening most patients do not take any hormonal medication that might suppress the cycle, but in other situations the influence of hormonal suppression on the detection of endometriosis has to be further evaluated.

This study focused on lesions visually diagnosed by two experienced gynecologist on laparoscopy instead of only pathological proven lesions as a reference standard. It was considered that direct visualisation with laparoscopy was a more realistic test to use as reference standard. Furthermore, the European Society of Human Reproduction and Embryology (ESHRE) stated in 2005 that visual inspection of the pelvis on laparoscopy can be used as the 'gold' standard investigation for a definite diagnosis of endometriosis [5].

In conclusion, using an optimized protocol, MR imaging seems reliable to detect all patients with endometriosis higher than stage I. Further studies have to be performed in larger patient populations and in patients with a low suspicion of the disease to confirm our findings.

## Acknowledgement

We wish to thank Roy S Dwarkasing, MD and Wim Hop, PhD for their valuable input.

## Disclosure

None of the authors have relationships with companies that may have a financial interest in the information contained in the manuscript.

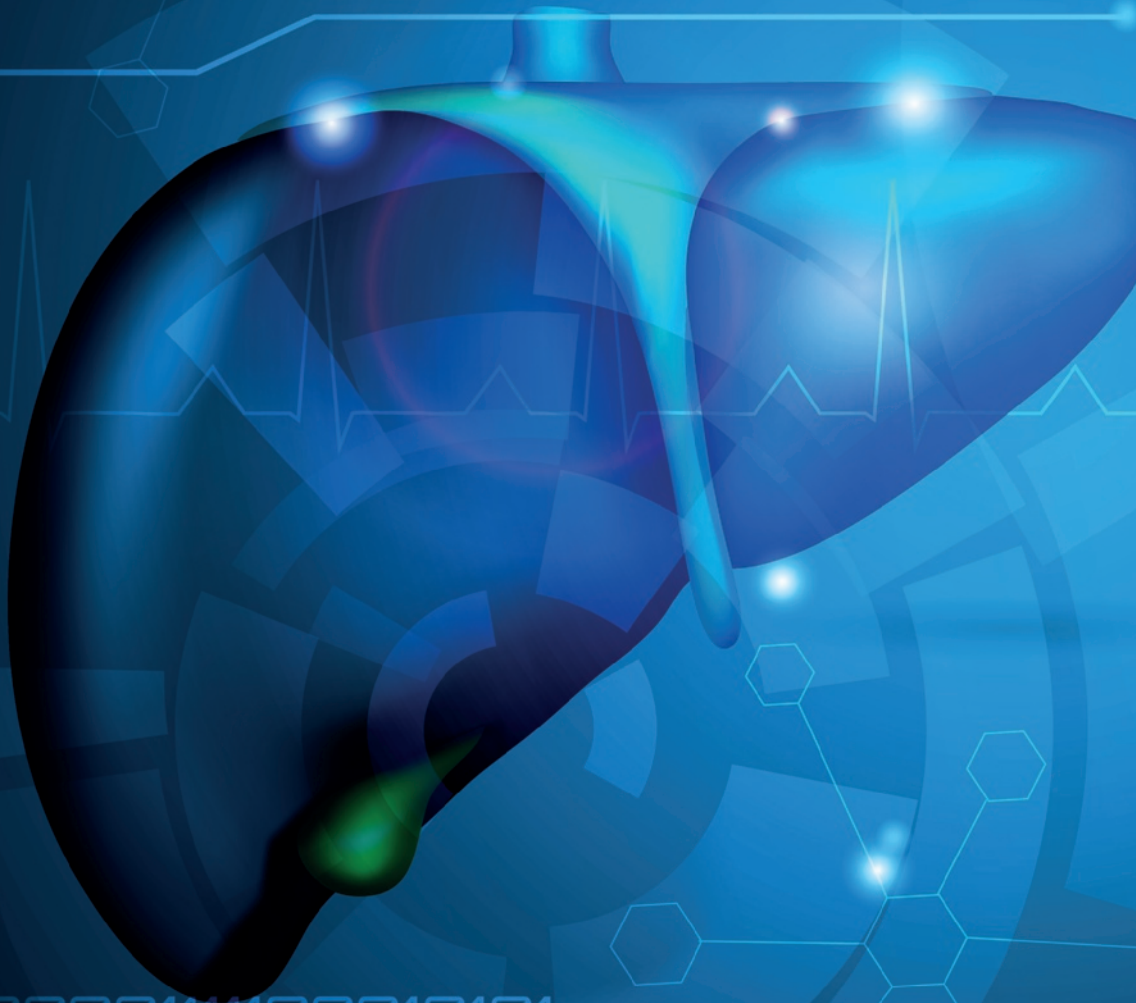
## References

1. Bulun SE. Endometriosis. *N Engl J Med* 2009;360(3):268-79.
2. Prevalence and anatomical distribution of endometriosis in women with selected gynaecological conditions: results from a multicentric Italian study. Gruppo italiano per lo studio dell'endometriosi. *Hum Reprod* 1994;9(6):1158-62.
3. Zondervan KT, Yudkin PL, Vessey MP, Dawes MG, Barlow DH, Kennedy SH. Patterns of diagnosis and referral in women consulting for chronic pelvic pain in UK primary care. *Br J Obstet Gynaecol* 1999;106(11):1156-61.
4. Ballard K, Lowton K, Wright J. What's the delay? A qualitative study of women's experiences of reaching a diagnosis of endometriosis. *Fertil Steril* 2006;86(5):1296-301.
5. Kennedy S, Bergqvist A, Chapron C, et al. ESHRE guideline for the diagnosis and treatment of endometriosis. *Hum Reprod* 2005;20(10):2698-704.
6. Bedaiwy MA, Falcone T, Sharma RK, et al. Prediction of endometriosis with serum and peritoneal fluid markers: a prospective controlled trial. *Hum Reprod* 2002;17(2):426-31.
7. Martinez S, Garrido N, Coperias JL, et al. Serum interleukin-6 levels are elevated in women with minimal-mild endometriosis. *Hum Reprod* 2007;22(3):836-42.
8. Othman Eel D, Hornung D, Salem HT, Khalifa EA, El-Metwally TH, Al-Hendy A. Serum cytokines as biomarkers for nonsurgical prediction of endometriosis. *Eur J Obstet Gynecol Reprod Biol* 2008;137(2):240-6.
9. D'Hooghe TM, Mihalyi AM, Simsa P, et al. Why we need a noninvasive diagnostic test for minimal to mild endometriosis with a high sensitivity. *Gynecol Obstet Invest* 2006;62(3):136-8.
10. Mihalyi A, Gevaert O, Kyama CM, et al. Non-invasive diagnosis of endometriosis based on a combined analysis of six plasma biomarkers. *Hum Reprod* 2010;25(3):654-64.
11. Seeber B, Sammel MD, Fan X, et al. Panel of markers can accurately predict endometriosis in a subset of patients. *Fertil Steril* 2008;89(5):1073-81.
12. Abrao MS, Goncalves MO, Ajossa S, Melis GB, Guerriero S. The sonographic diagnosis of deep endometriosis. *J Ultrasound Med* 2009;28(3):408-9; author reply 9-10.
13. Guerriero S, Alcazar JL, Pascual MA, et al. Diagnosis of the most frequent benign ovarian cysts: is ultrasonography accurate and reproducible? *J Womens Health (Larchmt)* 2009;18(4):519-27.
14. Hottat N, Larrousse C, Anaf V, et al. Endometriosis: contribution of 3.0-T pelvic MR imaging in preoperative assessment--initial results. *Radiology* 2009;253(1):126-34.
15. Stratton P, Winkel C, Premkumar A, et al. Diagnostic accuracy of laparoscopy, magnetic resonance imaging, and histopathologic examination for the detection of endometriosis. *Fertil Steril* 2003;79(5):1078-85.
16. Revised American Society for Reproductive Medicine classification of endometriosis: 1996. *Fertil Steril* 1997;67(5):817-21.
17. Bazot M, Darai E, Hourani R, et al. Deep pelvic endometriosis: MR imaging for diagnosis and prediction of extension of disease. *Radiology* 2004;232(2):379-89.
18. Kataoka ML, Togashi K, Yamaoka T, et al. Posterior cul-de-sac obliteration associated with endometriosis: MR imaging evaluation. *Radiology* 2005;234(3):815-23.
19. Park T. A comparison of the generalized estimating equation approach with the maximum likelihood approach for repeated measurements. *Stat Med* 1993;12(18):1723-32.
20. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33(1):159-74.
21. Catenacci M, Sastry S, Falcone T. Laparoscopic surgery for endometriosis. *Clin Obstet Gynecol* 2009;52(3):351-61.
22. Jacobson TZ, Duffy JM, Barlow D, Koninckx PR, Garry R. Laparoscopic surgery for pelvic pain associated with endometriosis. *Cochrane Database Syst Rev* 2009(4):CD001300.
23. Alliance EE. [www.endometriosis.org](http://www.endometriosis.org). In, 2006



## Part II

# Hepatocellular adenoma



000000111100010101

000000111100010101



# Chapter 5

## Hepatocellular adenomas: correlation of MR imaging findings with pathologic subtype classification

Sanne van Aalten

Maarten Thomeer

Turkan Terkivatan

Roy Dwarkasing

Joanna Verheij

Rob de Man

Jan IJzermans

## Abstract

**Purpose:** To investigate the correlation between magnetic resonance (MR) imaging findings and pathologic subtype classification of hepatocellular adenoma (HCA).

**Materials and Methods:** This retrospective study was approved by the institutional review board, and the requirement for informed consent was waived. MR imaging studies of 61 lesions (48 patients; median age, 36 years) were available and were independently reviewed by two radiologists. Consensus readings on all morphologic and signal-intensity imaging features were obtained. Previously, these lesions had been classified on the basis of pathologic findings and immunohistochemical analysis. Fisher exact and  $\chi^2$  tests were performed to compare the results between the different subtypes. A Bonferroni correction was applied to correct for multiple testing ( $\alpha < .0033$ ).

**Results:** MR imaging signs of diffuse intratumoral fat deposition were present in seven (78%) of nine liver-fatty acid binding protein (L-FABP)-negative HCAs compared with five (17%) of 29 inflammatory HCAs ( $P = .001$ ). Steatosis within the nontumoral liver was present in 11 (38%) of 29 inflammatory HCAs compared with none of the L-FABP-negative HCAs ( $P = .038$ ). A characteristic atoll sign was only seen in the inflammatory group ( $P = .027$ ). Presence of a typical vaguely defined type of scar was seen in five (71%) of seven  $\beta$ -catenin-positive HCAs ( $P = .003$ ). No specific MR imaging features were identified for the unclassified cases.

**Conclusion:** L-FABP-negative, inflammatory, and  $\beta$ -catenin-positive HCAs were related to MR imaging signs of diffuse intratumoral fat deposition, an atoll sign, and a typical vaguely defined scar, respectively. Since  $\beta$ -catenin-positive HCAs are considered premalignant, closer follow-up with MR imaging or resection may be preferred.

## Keywords

adenoma, liver, MRI, immunohistochemistry

## Introduction

Hepatocellular adenoma (HCA) is an uncommon benign tumor of the liver that occurs particularly in young and middle-aged women [1]. HCA can be complicated by life-threatening rupture and bleeding or undergo malignant transformation [2;3]. Management of HCA frequently requires cessation of oral contraceptives, intermittent follow-up with radiologic imaging, and a recommendation to avoid pregnancy [4;6]. Recently, the Bordeaux group established a new molecular and pathologic classification of HCA. They divided HCA into four different subgroups according to genotypic and phenotypic characteristics and clinical features [7-9]. The first group, comprising 35%–40% of the HCA cases, was defined by the presence of mutations in the T-cell factor 1 gene that inactivate the hepatocyte nuclear factor 1 $\alpha$ . Hepatocyte nuclear factor 1 $\alpha$ -inactivated HCAs lack expression of liver-fatty acid binding protein (L-FABP) in the tumor. As a result, the majority of these tumors are highly steatotic [10]. The second group, 10%–15% of cases, showed the presence of  $\beta$ -catenin-activating mutations.  $\beta$ -Catenin-activated lesions seem to have a higher risk of malignant transformation in patients with an HCA larger than 4–5 cm [7;8]. The third group, 50% of cases, showed inflammatory HCAs with serum amyloid A (SAA) and C-reactive protein (CRP) expression. Finally, less than 10% of HCAs were unclassified [8;10]. Glutamine synthetase (GS) is another useful marker in pathologic analysis of liver tumors. Strong, diffuse, and homogeneous GS staining has been shown in  $\beta$ -catenin-mutated HCA. This is in contrast to focal nodular hyperplasia (FNH), in which GS staining forms large hepatocyte areas that are organized in a characteristic maplike pattern, defined as positive GS staining in hepatocytes forming anastomosing areas, often surrounding venous structures [11]. As previously described by Laumonier et al [12], magnetic resonance (MR) imaging seems to be a useful tool for identifying the two major subtypes of HCA (ie, L-FABP-negative HCA and inflammatory HCA). The purpose of our study was to investigate the correlation between MR imaging findings and pathologic subtype classification of HCA.

## Advances in Knowledge

Liver-fatty acid binding protein (L-FABP)-negative hepatocellular adenoma (HCA) can be recognized with high reliability (seven [78%] of nine L-FABP-negative HCAs) as a diffuse decrease in signal intensity on opposed-phase MR images because of intratumoral fat deposition within the lesion.

In almost half of the inflammatory HCAs, we found a typical atoll sign on T2-weighted MR images, which was characterized by a hyperintense signal band in the periphery of the lesion and isointensity of the center of the lesion with respect of the surrounding liver.

MR imaging findings of a vaguely demarcated scar and poorly delimited high-signal-intensity areas on T2-weighted images seem to be related to  $\beta$ -catenin positivity.

## Materials and Methods

### Patients

This retrospective study was approved by the institutional ethical review board, and informed consent was waived.

### Implications for Patient Care

Steatotic HCAs smaller than 5 cm may be observed, and follow-up may be stopped when the lesion remains stable or shows reduction in size since we did not observe  $\beta$ catenin positivity in L-FABP- negative HCAs.

HCAs with a central scar and without other signs of focal nodular hyperplasia require a more aggressive approach because of the risk of malignant transformation.

Between January 2000 and February 2010, a cohort of 58 patients (71 lesions) who were radiologically diagnosed with an HCA were treated with surgical resection. The lesions were previously reviewed and classified on the basis of pathologic and immunohistochemical analysis [13]. Among these 58 patients, we identified 50 patients in whom liver MR imaging had been performed. Two patients were excluded from the study owing to a lack of adequate MR imaging pulse sequences to allow evaluation. A total of 48 patients (median age, 36 years) were included in the study. Among this group, a total of 61 lesions were surgically excised.

### Histopathologic and Immunohistochemical Analysis

Immunohistochemical staining was performed on paraffin sections representative of the tumor and included L-FABP (polyclonal antibody, 1:100 dilution; Bio-Connect, Huissen, the Netherlands), SAA (monoclonal mouse antibody, 1:200 dilution; Dako, Heverlee, Belgium), CRP (mouse monoclonal antibody, 1:1600 dilution; Bio-Connect), GS (monoclonal mouse antibody, 1:3200 dilution; BD Biosciences, Breda, the Netherlands), and  $\beta$ -catenin (monoclonal mouse antibody, 1:50 dilution; BD Biosciences), as previously described [13]. For each immunohistochemical stain, positive and negative controls were used for quality control. The immunohistochemical analysis allowed us to classify the adenomas into four subgroups, and subtype classification was straightforward in more than 90% of HCA cases. A lack of L-FABP expression, compatible with hepatocyte nuclear factor 1a mutation, was present in nine HCA lesions. Four lesions were positive for  $\beta$ -catenin without expression of CRP and SAA. Nuclear staining of  $\beta$ -catenin with a concomitant diffuse and homogeneous staining pattern for GS was observed in one of these four  $\beta$ -catenin-positive HCAs. All lesions in this subgroup showed cytologic atypia. The largest group consisted of 30 inflammatory HCAs with expression of both CRP and SAA. Additional  $\beta$ -catenin activation with a concomitant diffuse and homogeneous staining pattern for GS was demonstrated in three inflammatory HCAs. One inflammatory HCA demonstrated a heterogeneous staining pattern of GS without nuclear  $\beta$ -catenin staining positivity. Five HCAs remained unclassified because they were negative

for CRP, SAA,  $\beta$ -catenin, and GS, with a normal L-FABP staining pattern. Thirteen lesions that were morphologically compatible with FNH showed a characteristic maplike pattern of GS immunostaining; were negative for CRP, SAA, and  $\beta$ -catenin; and had normal L-FABP staining.

## MR Imaging Technique and Analysis

Forty-one patients were imaged at our institution. All these patients underwent MR imaging on one of two 1.5-T units (Philips Medical Systems, Best, the Netherlands; or Signa, GE, Milwaukee, Wis) with the same protocol. A fourchannel phased-array body coil was used. In each patient, a single-shot fast spin-echo sequence (repetition time msec/echo time msec, 832/80–120; flip angle, 90°) with varying echo times (short and long echo times of 80 and 120 msec, respectively), a fat-suppressed T2-weighted fast spinecho sequence (3000/80; flip angle, 90°), and T1-weighted in- and opposed-phase gradient-echo sequences (echo times, 4.6 and 2.3 msec, respectively; flip angle, 80°) were used. In addition, fat-suppressed dynamic gadolinium-enhanced T1-weighted imaging was performed in at least four phases (precontrast, arterial, portal, and delayed) after administration of an intravenous bolus of 30 mL of non-liver-specific gadolinium chelate (gadopentetate dimeglumine, Magnevist; Schering, Berlin, Germany). The optimal arterial phase was based on bolus tracking. The portal and delayed phases were acquired at 45 and 120 seconds, respectively, after the acquisition of the arterial phase. Delayed-phase threedimensional T1-weighted gradient-echo imaging was also performed at least 4 minutes after contrast material injection. A section thickness of 5–7 mm was used in all sequences. Seven patients were imaged at collaborating hospitals, and their studies were reviewed at our institution. Minimal prerequisites for inclusion of these studies were the availability of T1- and T2-weighted images of the whole liver. T1- and T2-weighted images were available for all seven MR imaging studies performed at collaborative hospitals. In six of the seven studies, in- and opposed-phase images and dynamic contrast material-enhanced images were available. In one MR imaging study obtained at a collaborating hospital, only uniphase contrast-enhanced images in the venous phase were available. T1- and T2-weighted, in- and opposedphase, and dynamic contrast-enhanced images were available for all MR imaging studies performed at our hospital.

Two readers (M.G.J.T. and R.S.D., both with 7 years experience in abdominal imaging) reviewed the MR images independently while blinded to the clinical history and pathologic diagnosis. Thereafter, consensus on all imaging features was obtained. For all lesions, the following image features were noted by the radiologist: number of nodules, presence of steatosis in the nontumoral liver (absence, mild, moderate, or severe), location of resected lesion according to the Couinaud numbering system [14], maximum size of resected lesion, contour (regular or lobulated), overall aspect on T2-weighted images (homogeneous or heterogeneous), presence of macroscopic hemorrhagic component (defined as focal T1-weighted hyperintense area), presence of necrotic or cystic component (defined as pronounced hyperintense signal on T2-weighted images), presence of central scar (defined as T2-hyperintense lines enhancing in the late venous phase), presence and percentage of region

of steatosis within the lesion (absence [0%], mild [ $< 33\%$ ], moderate [ $33\%–66\%$ ], or severe [ $> 66\%$ ]) (empirically), presence of tumor capsule (low signal intensity on T2-weighted images and delayed phase enhancement) or pseudocapsule (high signal intensity on T2-weighted images and delayed phase enhancement) [15], signal intensity on T1- and T2-weighted images (slightly hypointense, markedly hypointense, isointense, slightly hyperintense, or markedly hyperintense), predominant phase of enhancement (arterial, portal, or venous; homogeneous, heterogeneous, or two phases), and intensity of enhancement (slight, moderate, or marked). Finally, on the basis of the images evaluated, each radiologist was asked to make a diagnosis for each lesion detected. Differentiation between HCA and FNH was based on typical features for both lesions, as previously published [16–19].

## Statistical Analysis

Categorical variables were summarized as frequency and percentages. Categorized variables were compared with each other by using the  $\chi^2$  or Fisher exact test. All reported *P* values were based on two-sided tests of significance. A Bonferroni correction was applied to correct for multiple testing (15 categorical variables). A *P* value of less than .0033 was considered to indicate a significant difference. Because the  $\beta$ -catenin-activated and unclassified groups were quite small, the statistical analysis was mainly focused on the inflammatory and L-FABP-negative groups. Interobserver agreement was calculated by using Cohen  $\kappa$  statistics ( $\leq 0.20$ , poor agreement; 0.21–0.40, fair agreement; 0.41–0.60, moderate agreement; 0.61–0.80, good agreement; and 0.81–1.00, very good agreement) [20].

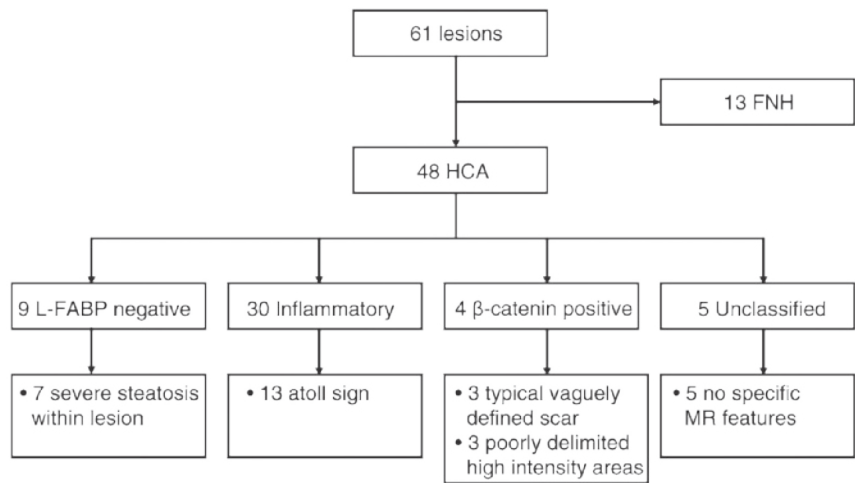
## Results

### Clinical and General Pathologic and Imaging Features

We identified HCA in 38 patients (48 lesions) and FNH in 10 patients (13 lesions) on the basis of morphologic criteria and immunohistochemical analysis (Figure 1). Forty-six (96%) of forty-eight patients were women (median age, 36 years; range, 15–64 years). The two male patients were both diagnosed with HCA (median age, 35 years; range, 31–39 years). The median age at diagnosis for all individuals was 36 years (range, 15–64 years). Oral contraceptive use was documented in 40 women. The duration of oral contraceptive use was known in 27 women: 24 women (89%) had been taking oral contraceptives for at least 5 years. Acute abdominal pain was the main reason for diagnostic imaging in 25 patients (52%), including bleeding of the lesion in six patients. Nonspecific symptoms were the reason for diagnostic imaging in seven cases, and the remaining cases were discovered as an incidental finding during radiologic examination of the abdomen for unrelated reasons. At MR imaging, 18 patients with HCA were considered to have a single lesion in the liver. Multiple adenomas (2–9 nodules) were found in 15 patients, and adenomatosis ( $\geq 10$  nodules) was present in five patients. In three patients, multiple resected HCAs were of the same subtype (one patient with three L-FABP-negative HCAs; two patients with inflammatory HCAs [two or three lesions]). Two patients had



one unclassified and one inflammatory HCA, and another patient had two  $\beta$  catenin-positive HCAs and one unclassified HCA.



**Figure 1.** Flowchart shows specific MR imaging features according to pathologic subtype classification of HCA.

At MR imaging (Table 1), 40 (85%) of 47 HCAs (one was not interpretable) were isointense to slightly hyperintense on T1-weighted images. On T2-weighted images, most HCAs were slightly hyperintense (37 of 48, 77%). No FNH lesions were hyperintense on T1-weighted images or contained intratumoral fat. On T2-weighted images, the signal was slightly hyperintense in eight (62%) of 13 FNH lesions. A central scar was seen Figure 1 in 10 (77%) of 13 FNH lesions, compared with 10 (21%) of 47 HCAs (one was doubtful) ( $P = .001$ ). Markedly intense homogeneous enhancement in the arterial phase was seen in eight (62%) of 13 FNH lesions, compared with 12 (26%) of 46 HCAs (two were not interpretable). On T2-weighted images, a characteristic hyperintense rimlike band in the periphery of the lesion (like an atoll) that was isointense to surrounding liver in the center of the lesion (like the surrounding sea) was seen in 13 (27%) of 42 (doubtful in six) of HCAs and in no FNH lesions. We termed this the atoll sign (Figures 2, 3). Interobserver agreement for the final diagnosis of each lesion was moderate ( $\kappa = 0.484$ ).

### MR Image Analysis by Subgroup

#### *L-FABP-negative HCAs*

Decreased signal intensity on T1-weighted opposedphase images owing to intratumoral fat deposition was present in seven (78%) of nine L-FABP-negative HCAs, measured as a region of severe and diffuse steatosis within the lesion (Table 2) (Figure 4). None of the other subtypes had diffuse steatosis within the lesion. Five (17%) of 29 inflammatory HCAs (one was

not interpretable) showed only focal steatosis within the lesion ( $P = .001$ ). All L-FABP-negative HCAs were slightly hyperintense on T2-weighted images.

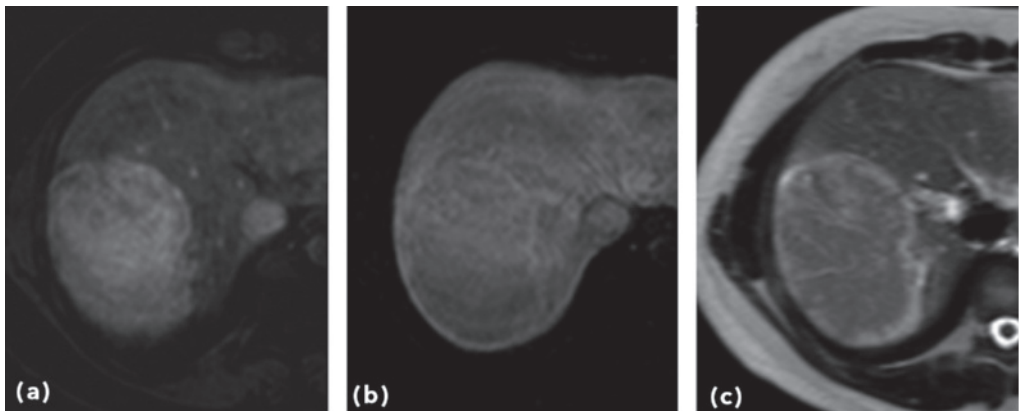
**Table 1.** MR Imaging Characteristics of HCA and FNH

Characteristic	HCA (n = 48)	FNH (n = 13)
Contour		
Regular	27 (56)	2 (15)
Lobulated	21 (44)	11 (85)
Aspect		
Homogeneous	17 (35)	6 (46)
Heterogeneous	31 (65)	7 (54)
Hemorrhagic component	10 (21)	0
Necrotic or cystic component		
Present	7 (15)	0
Doubtful	1	0
Central scar		
Present	10 (21) *	10 (77) *
Doubtful	1	0
Lesion steatosis <sup>1</sup>		
Absent	35 (74)	13 (100)
Mild	4 (9)	0
Moderate	1 (2)	0
Severe	7 (15)	0
Not interpretable	1	0
Capsule	0	0
Pseudocapsule	14 (29)	5 (38)
T1-weighted signal intensity		
Slightly hypointense	5 (11)	8 (62)
Markedly hypointense	1 (2)	5 (38)
Isointense	29 (62)	0
Slightly hyperintense	11 (23)	0
Markedly hyperintense	1 (2)	0
Not interpretable	1	0
T2-weighted signal intensity		
Slightly hypointense	3 (6)	2 (15)
Markedly hypointense	0	0
Isointense	7 (15)	3 (23)
Slightly hyperintense	37 (77)	8 (62)
Markedly hyperintense	1 (2)	0
Not interpretable	0	0

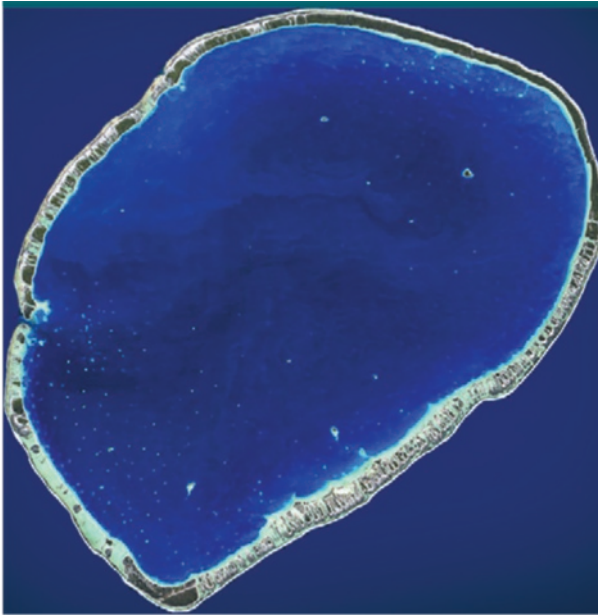
**Table 1.** Continued

Characteristic	HCA (n = 48)	FNH (n = 13)
Enhancement phase		
Arterial	47 (100)	13 (100)
Portal	0	0
Venous	0	0
Not interpretable	1	0
Enhancement		
Homogeneous	26 (57)	8 (62)
Heterogeneous	14 (30)	5 (38)
Two phases	6 (13)	0
Not interpretable	2	0
Intensity of enhancement		
Slight	5 (11)	0
Moderate	24 (51)	2 (15)
Marked	18 (38)	11 (85)
Not interpretable	1	0
Atoll sign		
Present	13 (27)	0
Doubtful	6 (13) <sup>1</sup>	(8)

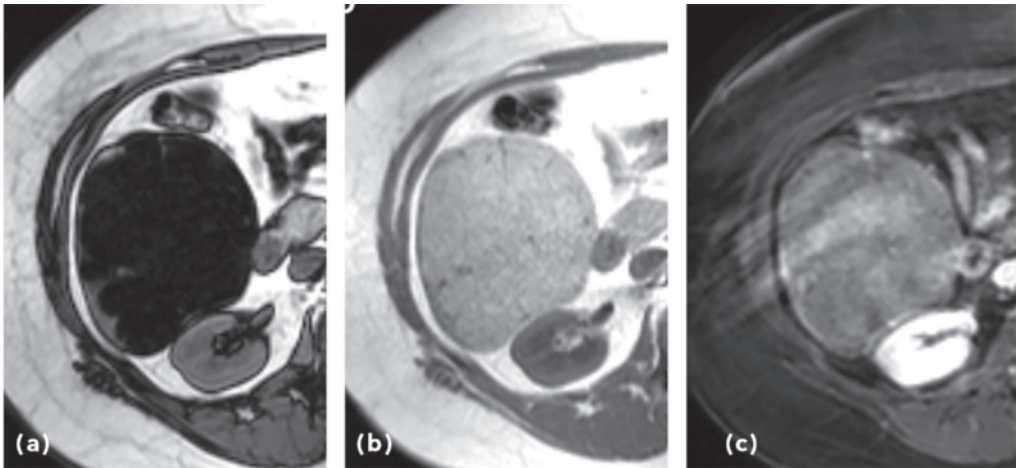
Note: Data are numbers of lesions, with percentages of lesions with interpretable results for that characteristic in parentheses. \*  $P = .001$  for difference between HCA and FNH. <sup>1</sup>Absent = 0%, mild = < 33%, moderate = 33%–66%, severe = > 66%.



**Figure 2.** (a) Arterial- and (b) late vascular-phase dynamic contrast-enhanced T1-weighted transverse MR images. (c) T2-weighted MR image shows characteristic atoll sign of an infl ammatory HCA in segment 6/7, which includes a T2-hyperintense signal band in the periphery of the lesion (like an atoll) with a center that is isointense to surrounding liver (like the surrounding sea). T2-hyperintense region typically enhances in the (b) late vascular phase, possibly corresponding to dilated sinusoids within infl ammatory HCA.



**Figure 3.** Satellite image of southern part of Tikehau Atoll. (Image courtesy of National Aeronautics and Space Administration Earth Observatory; <http://earthobservatory.nasa.gov/IOTD/view.php?id=39329> .)



**Figure 4.** Transverse (a) opposed-phase, (b) in-phase, and (c) contrast-enhanced venous-phase T1-weighted MR images show an L-FABP-negative HCA in segment 5/6, which shows a strong homogeneous drop in signal intensity on a and slight enhancement on c.

***β-catenin-positive HCAs***

Nuclear staining of β-catenin (without expression of CRP and SAA) was observed in four HCAs. On MR images, three (75%) of these nodules demonstrated a scar and poorly delimited high-signal-intensity areas (Figure 5). A pseudocapsule was present in three (75%) of the four β-catenin-positive HCAs, whereas none of these had a fibrous tumor capsule. One lesion had a hemorrhagic component.

***Inflammatory HCAs***

Twenty-four (80%) of 30 inflammatory HCAs appeared isointense to slightly hyperintense to the surrounding liver on T1-weighted images and appeared slightly hyperintense on T2-weighted images. Steatosis within the nontumoral liver was present in 11 (38%) of 29 inflammatory HCAs (one was not interpretable), whereas none of the L-FABP-negative HCAs showed steatosis within the nontumoral liver ( $P = .038$ ). The characteristic atoll sign was seen in 13 (43%) of 26 inflammatory HCAs, and a doubtful atoll sign was noted in four lesions. No clear cases with an atoll sign were seen in the L-FABP-negative group ( $P = .027$ ) or other subgroups. Focal positivity for β-catenin with a concomitant diffuse and homogeneous staining pattern of GS was seen in three inflammatory HCAs. None of these β-catenin-positive inflammatory HCAs showed steatosis within the nontumoral in 10 (77%) of 13 FNH lesions, compared liver, and only one lesion showed mild steatosis within the lesion. These lesions appeared heterogeneous on images obtained with each sequence. These lesions appeared isointense to slightly hyperintense to the surrounding liver on T1-weighted images and appeared slightly hyperintense on T2-weighted images. In accordance with β-catenin-positive HCAs without expression of CRP and SAA (see above), two β-catenin-positive inflammatory HCAs showed a central scar. None of the β-catenin-positive inflammatory HCAs had a hemorrhagic component. One inflammatory HCA demonstrated a heterogeneous staining pattern of GS without nuclear β-catenin staining. This lesion was heterogeneous on all MR images. On T1- and T2-weighted images, this lesion appeared slightly hyperintense to the surrounding liver. This lesion also had a central scar.

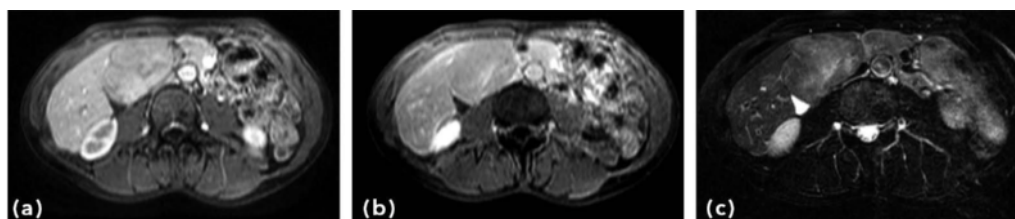
***Unclassified HCAs***

A small group of five HCAs remained unclassified with immunohistochemical analysis. In this group, no characteristic features were seen on MR images. None of these lesions showed steatosis within the lesion. Hemorrhagic components were seen in three (60%) of the five unclassified lesions.

***Central scar in β-catenin- and GSpositive lesions***

Ten (77%) of 13 FNH lesions had a typical scar revealing hyperintense lines on T2-weighted images, which enhanced in the portovenous phase. Ten (21%) of 47 HCAs had a central scar, and one HCA had a central scar that both readers interpreted as doubtful. Three of four β-catenin-positive HCAs (without expression of CRP and SAA) showed poorly delimited high-signal-intensity areas on T2-weighted images as hyperintense scar regions, which enhanced late

during the portovenous phase. These regions were revealed as a spotty less-demarcated pattern of enhancement (Figure 5). Three inflammatory HCAs showed a central scar, which was revealed as hypointense thin lines on T2-weighted images. Two inflammatory HCAs with nuclear positivity for  $\beta$ -catenin and a homogeneous GS staining pattern also showed a scar. Signs of a scar were present in one L-FABP-negative HCA and in one unclassified HCA.  $\beta$ -Catenin positivity in HCA and GS positivity in HCA and FNH were significantly associated with the presence of a central scar. Presence of a central scar was seen in five (71%) of seven  $\beta$ -catenin-positive HCAs compared with five (13%) of 40  $\beta$ -catenin-negative HCAs ( $P = .003$ ) (one was doubtful). HCA and FNH showed the presence of a central scar in 14 (78%) of 18 GS-positive lesions compared with six (14%) of 42 GS-negative lesions (one was doubtful) ( $P < .001$ ).



**Figure 5.** Transverse (a) arterial- and (b) late vascular-phase T1-weighted dynamic contrast-enhanced and (c) T2-weighted fat-saturated MR images show a  $\beta$ -catenin-positive HCA located in segment 4, with poorly delineated areas of increased signal intensity and a hyperintense scar on c. Lesion enhances irregularly on b.

The final diagnosis of the pathologist was compared with the final consensus interpretation of the two abdominal radiologists. In six cases, the diagnosis differed: Two of 48 HCAs were mistaken for FNH, and four of 13 FNH lesions were mistaken for HCA. Both false-negative HCAs showed presence of scar regions. Two of the four missed FNH lesions showed no scar, and one missed FNH was wrongly diagnosed owing to multiple smaller liver lesions reminiscent of multiple adenomas. All false-negative FNH lesions showed a homogeneous isointense or slightly hyperintense signal on T2-weighted images.

**Table 2.** MR Imaging Characteristics of HCA by Subtype

Characteristic	L-FABP-negative (n = 9)	Inflammatory (n = 30)	$\beta$ -Catenin-positive (n = 4)	Unclassified (n = 5)
Size (mm) *	75 (28–124)	77 (15–184)	51 (25–80)	51 (42–110)
Nontumoral liver steatosis				
Absent	9 (100)	18 (62)	4 (100)	3 (60)
Mild	0	2 (7)	0	1 (20)
Moderate	0	2 (7)	0	0
Severe	0	7 (24)	0	1 (20)
Not interpretable	0	1	0	0
Contour				
Regular	4 (44)	18 (60)	3 (75)	2 (40)
Lobulated	5 (56)	12 (40)	1 (25)	3 (60)
Aspect				
Homogeneous	4 (44)	10 (33)	2 (50)	1 (20)
Heterogeneous	5 (56)	20 (67)	2 (50)	4 (80)
Hemorrhagic component	2 (22)	4 (13)	1 (25)	3 (60)
Necrotic or cystic component				
Present	2 (22)	2 (7)	2 (50)	1 (20)
Doubtful	0	1	0	0
Central scar				
Present	1 (11)	5 (17)	3 (75)	1 (20)
Doubtful	0	1	0	0
Lesion steatosis <sup>1</sup>				
Absent	2 (22)	24 (83)	4 (100)	5 (100)
Mild	0	4 (14)	0	0
Moderate	0	1 (3)	0	0
Severe	7 (78)	0	0	0
Not interpretable	0	1	0	0
Capsule	0	0	0	0
Pseudocapsule	2 (22)	7 (23)	3 (75)	2 (40)
T1-weighted signal intensity				
Slightly hypointense	1 (13)	4 (13)	0	0
Markedly hypointense	0	1 (3)	0	0
Isointense	6 (75)	16 (53)	4 (100)	3 (60)
Slightly hyperintense	1 (13)	8 (27)	0	2 (40)
Markedly hyperintense	0	1 (3)	0	0
Not interpretable	1	0	0	0

**Table 2.** Continued

Characteristic	L-FABP-negative (n = 9)	Infl ammatory (n = 30)	β-Catenin-positive (n = 4)	Unclassified (n = 5)
T2-weighted signal intensity				
Slightly hypointense	0	2 (7)	0	1 (20)
Markedly hypointense	0	0	0	0
Isointense	0	3 (10)	2 (50)	2 (40)
Slightly hyperintense	9 (100)	24 (80)	2 (50)	2 (40)
Markedly hyperintense	0	1 (3)	0	0
Not interpretable	0	0	0	0
Enhancement phase				
Arterial	9 (100)	29 (100)	4 (100)	5 (100)
Portal	0	0	0	0
Venous	0	0	0	0
Not interpretable	0	1	0	0
Enhancement				
Homogeneous	6 (67)	16 (57)	2 (50)	2 (40)
Heterogeneous	3 (33)	8 (29)	2 (50)	1 (20)
Two phases	0	4 (14)	0	2 (40)
Not interpretable	0	2	0	0
Intensity enhancement				
Slight	3 (33)	1 (3)	1 (25)	4 (80)
Moderate	3 (33)	14 (48)	3 (75)	0
Marked	3 (33)	14 (48)	0	1 (20)
Not interpretable	0	1	0	0
Atoll sign				
Present	0	13 (43)	0	0
Doubtful	2 (22)	4 (13)	0	0

Note.—Unless otherwise specified, data are numbers of lesions, with percentages of lesions with interpretable results for that characteristic in parentheses. \* Data are medians, with ranges in parentheses. <sup>1</sup>Absent = 0%; mild = < 33%, moderate = 33%–66%; severe = > 66%. *P* = .001 for L-FABP-negative vs inflammatory HCAs.

## Discussion

The main goal of our study was to investigate the correlation between MR imaging findings and pathologic subtype classification since state-of-the-art MR imaging provides the most comprehensive and noninvasive imaging work-up of patients suspected of having HCAs [15]. We were able to confirm specific MR imaging features that can be used to identify different



subgroups of HCA, especially inflammatory and L-FABP- negative HCAs, as previously described by Laumonier et al [12].

L-FABP-negative HCA can be recognized with high reliability and is seen as a diffuse decrease in signal intensity on opposed-phase images because of the presence of intratumoral fat in the lesion. MR imaging was used to detect L-FABP- negative HCA in 78% of L-FABP-negative adenomas in which marked intratumoral steatosis was present. Only 17% of inflammatory HCAs showed steatosis within the lesion ( $P = .001$ ); however, this subgroup showed only mild to moderate steatosis, which was distributed focally. All L-FABP-negative HCAs showed slightly hyperintense signal on T2-weighted images, possibly owing to the intratumoral T2 effect of fat (J-coupling effect) [21].

In almost half of the inflammatory HCAs, we found a typical atoll sign on MR images, which is characterized by a hyperintense signal band in the periphery of the lesion (like an atoll) and isointensity of the center of the lesion with respect to the surrounding liver (like the surrounding sea) on T2-weighted images. No clear cases with an atoll sign were seen in the L-FABP-negative or other subgroups; however, the difference did not reach significance since a Bonferroni correction was applied to correct for multiple testing. In 10 (77%) of 13 inflammatory HCAs with a characteristic atoll sign, T2-hyperintense islands were visible in the center portion of the lesion. The T2-hyperintense regions (signal band and center islands) typically enhance in the late vascular phase. This characteristic atoll sign is possibly due to sinusoidal dilatation within inflammatory HCA. These areas of sinusoidal dilatation enhance in the venous phase, whereas the rest of the tumor enhances in the arterial phase. Those areas are hyperintense on T2-weighted images, which is probably reflective of the high water content in these bloodfilled regions.

The presence of a vaguely demarcated scar in HCA was significantly associated with nuclear staining of  $\beta$ -catenin ( $P = .003$ ). At microscopic examination of the available slides from the tumor, two of five  $\beta$ -catenin-positive HCAs with presence of a vaguely demarcated scar showed pathologic areas and septa of fibrosis. The remaining  $\beta$ -catenin-positive HCAs with presence of a scar did not show areas of fibrosis, which may be owing to sampling error because of the large size (29, 70, and 100 mm) and limited sampling of these lesions. Presence of a scar in HCA and FNH was significantly associated with a diffuse and homogeneous GS staining pattern ( $P < .001$ ). However, three lesions presented a central scar without positive  $\beta$ -catenin and GS findings. All  $\beta$ -catenin-positive HCAs showed features of slight to moderate enhancement compared with FNH, which showed marked enhancement in most cases. We did not observe  $\beta$ -catenin or GS positivity in L-FABP- negative lesions. Therefore, by combining the atoll sign with the presence of a scar, we may be able to differentiate between  $\beta$ -catenin-positive inflammatory HCA and  $\beta$ -catenin-negative inflammatory HCA. Typical poorly delimited high-signal-intensity areas on T2-weighted images may be another characteristic MR imaging

feature observed in  $\beta$ -catenin-positive HCAs, since this was also described by Laumonier et al [12]. No specific MR imaging features were identified for the unclassified group of HCA.

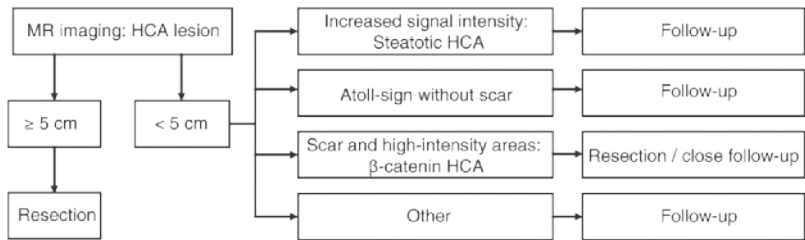
In most cases in our study, a clear diagnosis of HCA or FNH could be made. However, two (4%) of 48 HCAs were mistaken for FNH, and four (31%) of 13 FNH lesions were mistaken for HCA. The high percentage of FNH lesions that were mistaken for HCAs may be owing to a verification bias since the FNH lesions included were diagnosed as HCAs at the time of resection. It is likely that more atypical FNH lesions were included in our study. The verification bias may also be a reason for the low percentage (62%) of FNH lesions showing intense homogeneous enhancement. The moderate interobserver agreement, mainly owing to the verification bias of atypical FNH lesions, may suggest that the identification of imaging features of benign hepatic lesions such as HCA and FNH may be quite challenging at times.

The question is: How can we differentiate HCA from FNH with more certainty? It is probably that accurate differentiation requires the combination of different features. Hyperintensity on both in- and opposed-phase T1-weighted images (owing to intratumoral blood degradation products) is often seen in HCA and has rarely been noted in FNH lesions [22]. Intratumoral fat deposition is a typical feature in HCA and is seldom seen in FNH. While slight enhancement in the arterial phase can be seen in HCA, typical FNH lesions have more pronounced enhancement. Moreover, the presence of a scar in the lesion is not solely found in FNH lesions but may also be seen in HCAs. In all HCA cases in our study where there was a misdiagnosis (probably owing to the internal scar), the slight degree of enhancement of the lesion should have alerted us that the lesion could be HCA instead of FNH.

We suggest classifying HCA by using state-of-the-art MR imaging and following the Bordeaux criteria. Specific subgroups of patients with HCA could be counseled appropriately. In women with a clear diagnosis of HCA lesions at least 5 cm in diameter that do not show adequate regression after discontinuation of oral contraceptives, we recommend surgical resection. Steatotic HCAs smaller than 5 cm may be observed, and follow-up may be stopped when the lesion remains stable or shows reduction in size after 6 months of follow-up, since we did not observe  $\beta$ -catenin positivity in L-FABP-negative HCAs. HCAs that present with a central scar and without other signs of FNH require a more aggressive approach because of the risk of malignant transformation. Close follow-up or resection may be advised in these cases (Figure 6). No published data are available on the success rate of using immunohistochemical markers to classify subgroups at biopsy. The interpretation of  $\beta$ -catenin staining on biopsies may be difficult because of the heterogeneous and sometimes focal nuclear staining pattern, and our previous experience suggests that sampling error will remain a problem in this situation [13].

The major limitation of our study was the small number of  $\beta$ -catenin-positive HCAs (four lesions of the  $\beta$ -catenin subtype and three  $\beta$ -catenin-positive inflammatory lesions) we were able to include. The identification of inflammatory  $\beta$ -catenin-positive HCA remains a crucial

issue owing to the risk of malignant transformation. Therefore, the poorly delimited high-signal-intensity areas observed in  $\beta$ -catenin-positive HCAs deserve additional investigation. Multicenter studies will be required to improve the identification of HCA subgroups and to investigate the consequences for clinical management.



**Figure 6.** Flowchart shows suggested surgical management of HCA on the basis of MR imaging findings.

MR imaging may improve the subtype classification of HCA, especially in cases of inflammatory and L-FABP-negative HCAs. MR imaging findings of a vaguely demarcated scar and poorly delimited high-signal-intensity areas on T2-weighted images in HCA seem to be related to  $\beta$ -catenin positivity. Since  $\beta$ -catenin-positive lesions seem to have a higher risk of malignant transformation, surgical resection or close follow-up is advocated in patients presenting with these lesions.

## References

1. van der Sluis FJ, Bosch JL, Terkivatan T, de Man RA, IJzermans JN, Hunink MG. Hepatocellular adenoma: cost-effectiveness of different treatment strategies. *Radiology* 2009; 252(3):737-746.
2. Stoot JH, Coelen RJ, De Jong MC, Dejong CH. Malignant transformation of hepatocellular adenomas into hepatocellular carcinomas: a systematic review including more than 1600 adenoma cases. *HPB (Oxford)* 2010; 12(8):509-522.
3. Terkivatan T, de Wilt JH, de Man RA, van Rijn RR, Tilanus HW, IJzermans JN. Treatment of ruptured hepatocellular adenoma. *Br J Surg* 2001; 88 (2):207-209.
4. van Aalten SM, Terkivatan T, de Man RA, et al. Diagnosis and treatment of hepatocellular adenoma in the Netherlands: similarities and differences. *Dig Surg* 2010; 27 (1):61-67.
5. Choi BY, Nguyen MH. The diagnosis and management of benign hepatic tumors. *J Clin Gastroenterol* 2005; 39 (5):401-412.
6. Lin H, van den Esschert J, Liu C, van Gulik TM. Systematic review of hepatocellular adenoma in China and other regions. *J Gastroenterol Hepatol* 2011; 26 (1):28-35.
7. Bioulac-Sage P, Rebouissou S, Thomas C, et al. Hepatocellular adenoma subtype classification using molecular markers and immunohistochemistry. *Hepatology* 2007; 46 (3):740-748.
8. Zucman-Rossi J, Jeannot E, Nhieu JT, et al. Genotype-phenotype correlation in hepatocellular adenoma: new classification and relationship with HCC. *Hepatology* 2006; 43 (3):515-524.
9. Bioulac-Sage P, Balabaud C, Zucman-Rossi J. Subtype classification of hepatocellular adenoma. *Dig Surg* 2010; 27 (1):39-45.
10. Bioulac-Sage P, Laumonier H, Couchy G, et al. Hepatocellular adenoma management and phenotypic classification: the Bordeaux experience. *Hepatology* 2009; 50 (2):481-489.
11. Bioulac-Sage P, Laumonier H, Rullier A, et al. Over-expression of glutamine synthetase in focal nodular hyperplasia: a novel easy diagnostic tool in surgical pathology. *Liver Int* 2009; 29 (3):459-465.
12. Laumonier H, Bioulac-Sage P, Laurent C, Zucman-Rossi J, Balabaud C, Trillaud H. Hepatocellular adenomas: magnetic resonance imaging features as a function of molecular pathological classification. *Hepatology* 2008; 48 (3):808-818.
13. van Aalten SM, Verheij J, Terkivatan T, Dwarkasing RS, de Man RA, IJzermans JN. Validation of a liver adenoma classification system in a tertiary referral centre: implications for clinical practice. *J Hepatol* 2011; 55 (1):120-125.
14. Couinaud C. Le foie. In: *Études anatomiques et chirurgicales*. Paris, France: Masson, 1957; 74-75.
15. Hussain SM, van den Bos IC, Dwarkasing RS, Kuiper JW, den Hollander J. Hepatocellular adenoma: findings at state-of-the-art magnetic resonance imaging, ultrasound, computed tomography and pathologic analysis. *Eur Radiol* 2006; 16 (9):1873-1886.
16. Arrivé L, Fléjou JF, Vilgrain V, et al. Hepatic adenoma: MR findings in 51 pathologically proven lesions. *Radiology* 1994; 193 (2):507-512.
17. Grazioli L, Federle MP, Brancatelli G, Ichikawa T, Olivetti L, Blachar A. Hepatic adenomas: imaging and pathologic findings. *RadioGraphics* 2001; 21 (4):877-892; discussion 892-894.
18. Mortelé KJ, Praet M, Van Vlierberghe H, de Hemptinne B, Zou K, Ros PR. Focal nodular hyperplasia of the liver: detection and characterization with plain and dynamic-enhanced MRI. *Abdom Imaging* 2002; 27 (6):700-707.
19. van den Bos IC, Hussain SM, de Man RA, Zondervan PE, IJzermans JN, Krestin GP. Primary hepatocellular lesions: imaging findings on state-of-the-art magnetic resonance imaging, with pathologic correlation. *Curr Probl Diagn Radiol* 2008; 37 (3):104-114.
20. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33 (1):159-174.
21. Henkelman RM, Hardy PA, Bishop JE, Poon CS, Plewes DB. Why fat is bright in RARE and fast spin-echo imaging. *J Magn Reson Imaging* 1992; 2 (5):533-540.
22. Braga L, Armao D, Semelka RC. Liver. In: Semelka RC, ed. *Abdominal-pelvic MRI*. 2nd ed. Hoboken, NJ: Wiley, 2006; 47-445.

# Chapter 6

## MRI features of inflammatory hepatocellular adenomas on hepatocyte phase imaging with liver-specific contrast agents

Maarten G Thomeer

Francois E Willemsen

Katharina K Biermann

H Addouli

Rob A de Man

Jan J IJzermans

Roy S Dwarkasing

## Abstract

**Purpose:** To evaluate the presentation of inflammatory hepatocellular adenomas (HCAs) on hepatocyte phase magnetic resonance imaging (MRI) .

**Materials and Methods:** We retrospectively reviewed the MRI features of histologically proven HCAs on hepatocyte phase imaging. 21 lesions (17 with inflammatory subtype) were scanned with gadobenate dimeglumine. Signal intensities of the lesions were assessed in the hepatocyte phase and on the T1-weighted sequences before contrast.

**Results:** After gadobenate dimeglumine injection, 71% (12/17) of the inflammatory HCAs showed areas of iso- or hyperintensity to the surrounding liver in the hepatocyte phase. In 82% (10/12) of the iso- or hyperintense lesions, this was found over more than 75% of the lesion surface. None of the non-inflammatory HCAs showed areas of iso- or hyperintensity to the surrounding liver in the hepatocyte phase.

From these 12, seven were hyperintense before contrast due to liver steatosis, two due to intrinsic hyperintensity (on the in-phase sequence) and three were isointense.

**Conclusion:** In contrast to non-inflammatory HCAs, inflammatory HCAs can show areas of iso- to hyperintensity to the surrounding liver in the hepatocyte phase; therefore other typical imaging features should also be used to distinguish between HCAs and FNHs.

## Keywords

Hepatocellular adenoma, Gadobenate dimeglumine, Gadoxetate disodium, MRI

## Introduction

Hepatocellular adenoma (HCA) and focal nodular hyperplasia (FNH) are benign focal lesions of the liver that occur mainly in women [1]. In contrast to focal nodular hyperplasia (FNH) of the liver, which is a uncomplicated and solely benign entity, HCA can involve complications such as malignant degeneration or life-threatening bleeding [2]. Once a HCA has been diagnosed, resection may be warranted if the tumor exceeds 5 cm in diameter, and follow-up advice including ending contraceptive use and the discouragement of pregnancy should be considered in cases with smaller lesions [1;3].

As routine histological biopsy of suspected HCA lesions may produce nonspecific findings [4] and can induce bleeding, HCAs are primarily diagnosed by non-invasive imaging techniques. Most HCAs can be easily differentiated from FNHs based on specific magnetic resonance imaging (MRI) characteristics. HCAs more commonly exhibit high signal intensity on T1-weighted sequences, more commonly show signs of fatty infiltration, show less intense homogeneous enhancement and less typically show scar formations (no spoke wheel configuration) [1;5]. Differentiation of these respective lesions can be challenging for a less experienced reader or if some typical features are not clearly present [6]. In difficult cases, MRI with a hepato-specific contrast agent may help in the further differentiation of these entities. Gadobenate dimeglumine and gadoxetate disodium are currently the two main hepato-specific contrast agents; they are taken up by hepatocytes through for instance OATP transporter channels and excreted in the biliary system [7]. Lesions such as metastases and HCAs normally do not harbour this OATP activity and do not contain bile ducts, therefore they generally appear hypointense in the hepatocyte phase, compared to the liver. In contrast, FNHs and some hepatocellular carcinomas are composed of functioning hepatocytes and bile ducts [8;9] and subsequently show iso- to hyperintense areas in the hepatocyte phase, relative to the liver. According to recent publication this can be categorized in five different patterns [10].

Based solely on this finding, the differentiation of HCAs and FNHs appears straightforward and this was apparently confirmed by the results of a multicenter study in which the authors reported that all 107 diagnosed HCAs appeared to be hypointense on hepatocyte imaging [11;12].

However, as exception one subtype of HCAs - the inflammatory subtype, including telangiectatic HCA – can contain histopathologically-evident bile ducts [12], one would expect at least a proportion of these lesions to be relatively iso- or hyperintense on hepatocyte imaging.

In order to determine if this subtype of HCAs indeed show areas iso- or hyperintensity on hepatocyte imaging, we retrospectively studied our MRI scans to analyze the behavior of gadobenate dimeglumine on inflammatory HCAs, both relative to the liver in the hepatocyte

phase and as an absolute uptake in comparison to pre-contrast images. Since the prevalence of MRI scans performed with gadoxetate disodium is low in our retrospective database we looked at the behavior of gadoxetate disodium only with a limited approach.

## Materials and Methods

### Study population

The appropriate institutional review board approved this retrospective study and informed consent was waived. We reviewed the MRI data of all patients who underwent MRI for a primary liver lesion between January 2005 and July 2012.

Gadobenate dimeglumine was used as the intravenous contrast agent at our department of Radiology, and gadoxetate disodium was used in other institutions. The latter images were reviewed at our department, a part of a tertiary referral institution, as a second opinion in clinical management.

The inclusion criteria for patients were: MRI enhanced with gadobenate dimeglumine or gadoxetate disodium, biopsy or resection of a liver lesion and a final diagnosis of the lesion as adenoma. This latter was based on immunohistochemistry analysis.

### Histopathological analysis

Immunohistochemistry was performed on representative paraffin tumor sections for all resected lesions, or on percutaneous biopsy specimens. Immunostainings included liver fatty acid binding protein (L-FABP) (polyclonal antibody, 1:100 dilution, Bio-Connect), serum amyloid A (SAA) (monoclonal mouse antibody, 1:200 dilution, Dako), C-reactive protein (CRP) (mouse monoclonal antibody, 1:1600 dilution, Bio-Connect), glutamate synthetase (GS) (monoclonal mouse antibody, 1:3200 dilution, BD Biosciences) and  $\beta$ -catenin (monoclonal mouse antibody, 1:50 dilution, BD Biosciences) as previously described [13]. Lesions were diagnosed as inflammatory HCAs if there was clear expression of both CRP and SAA. Ductular proliferation was visible in all inflammatory HCAs, but not in other lesions (Figure 1).

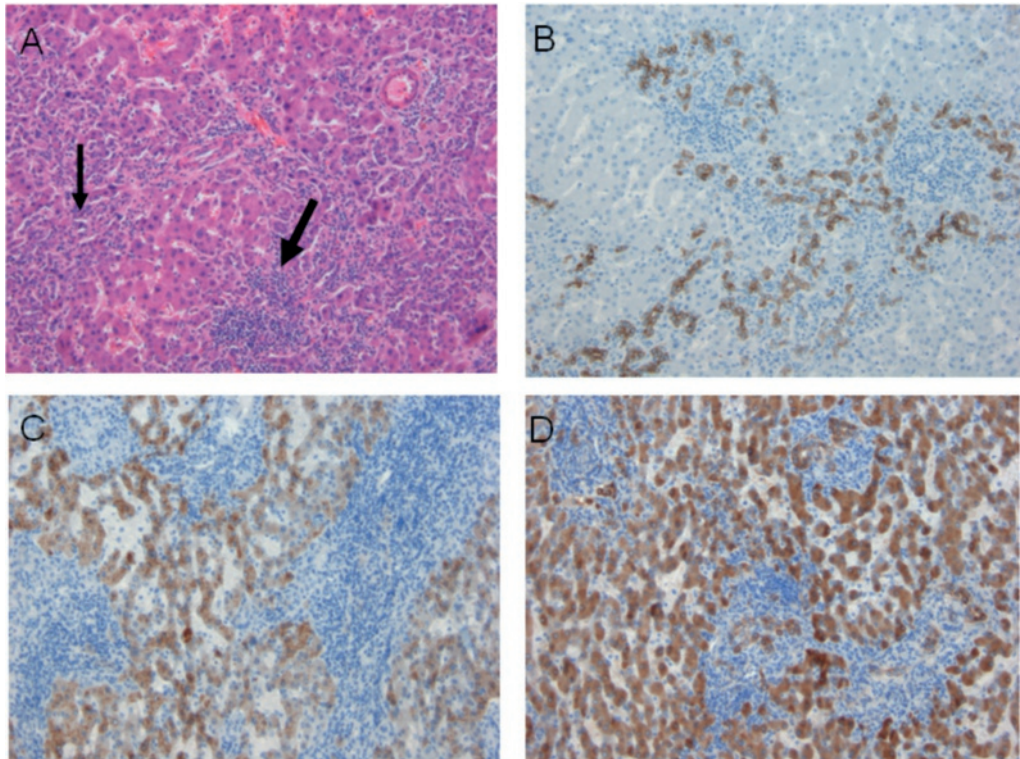
### Imaging investigation

All patients included at our department underwent MRI using a 1.5 T unit (Philips Medical Systems, Best, The Netherlands or General Electrics, Signa, Milwaukee, WI, USA; using an identical protocol). A four-channel, phased-array body coil was used. Scans were started with a single-shot fast spin echo (SSFSE, slice thickness = 7 mm; repetition time/echo time (TR/TE) (msec) = 832/80–120; flip angle = 90°), fat-suppressed T2W fast spin echo (FSE) (5–8 mm, 6315/90–93, flip angle = 90°), and T1-weighted in- and opposed-phase gradient-echo (GRE) sequences (7 mm, shortest/4.6 and 2.3, respectively; flip angle = 80°). Fat-suppressed, dynamic contrast-enhanced T1-weighted GRE imaging (4–5 mm, 2.7–3.5/1.2; flip



angle =  $12^\circ$ ) was performed in at least four phases (precontrast, arterial, portal, and delayed), following administration of an intravenous bolus (2–2.5 mL/sec) of gadobenate dimeglumine (Multihance, Bracco Imaging, Milan, Italy) at a dose of 0.05 mmol per kilogram body weight. The optimal arterial phase was based on bolus tracking. Finally, the same scan was repeated during a later hepatocyte phase at 1 hour after injection.

Patients scanned elsewhere and reviewed at our institution had the following minimal criteria: T2-weighted SSFSE (5 mm, 390–2976/57–80; flip angle =  $90^\circ$ ), T1-weighted, fat-saturated GRE sequence (5–8 mm, 3.5–4.2/1.6–2.1; flip angle =  $10$ – $15^\circ$ ) before and 20 minutes after intravenous administration of gadoxetate disodium (Primovist, Bayer Healthcare, Berlin, Germany) (U.S. brand name Eovist) at a dose of 0.025 mmol per kilogram body weight. The dynamic phase and precontrast chemical shift images were not reviewed since they were not used for study purposes.



**Figure 1. A-D:** example of a inflammatory adenoma (X100 magnification). **A:** HE slide (**A**) shows a mixture of neoplastic hepatocytes and ductular reaction with inflammatory infiltrate (arrow), also visible in **B** by CK7 immunohistochemistry (brownish color corresponds to ductular proliferation). **C, D:** the neoplastic hepatocytes are positive for serum amyloid A (**C**) and C-reactive protein (**D**) on immunohistochemistry.

## Image evaluation

Features of the extracted lesions diagnosed as inflammatory HCA were separately qualitatively analysed by 2 readers (MT and FW): these features were the signal intensity of the lesion (hypointense, isointense, hyperintense) in comparison to the surrounding liver on T1-weighted delayed phase fat-saturated sequences, and the percentage of lesion surface that was iso- or hyperintense to the liver (0%, 1–25%, 26–50%, 51–75%, 76–99% or 100%).

Finally, intensity of the lesions compared to the liver were analyzed on the precontrast T1-weighted images.

## Statistical analysis

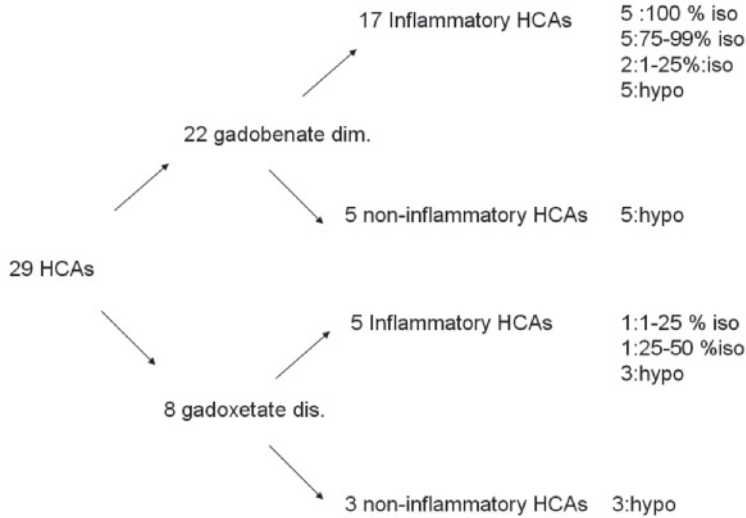
Interobserver agreement was assessed using the kappa value ( $\kappa$ ). A  $\kappa \leq 0.20$  was interpreted as slight agreement; 0.21–0.40, fair agreement; 0.41–0.60, moderate agreement; 0.61–0.80, substantial agreement; and  $\geq 0.81$ , almost perfect agreement [14].

## Results

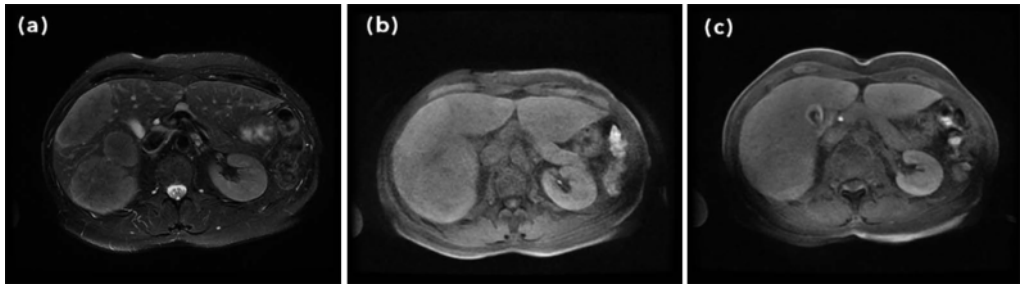
Twenty-nine proven HCAs were derived from 22 patients (Figure 2); 21 lesions (16 with inflammatory subtype) were scanned with gadobenate dimeglumine and 7 (4 with inflammatory subtype) with gadoxetate disodium (Figure 2). One supplementary inflammatory HCA was scanned both with gadobenate dimeglumine and gadoxetate disodium. Histological diagnosis and subtype classification was based on either resection (24 lesions), or on biopsy (5 lesions).

The median age of patients was 38 years (range 26–61 years). The median size of the lesions was 6 cm (range 2–17 cm).

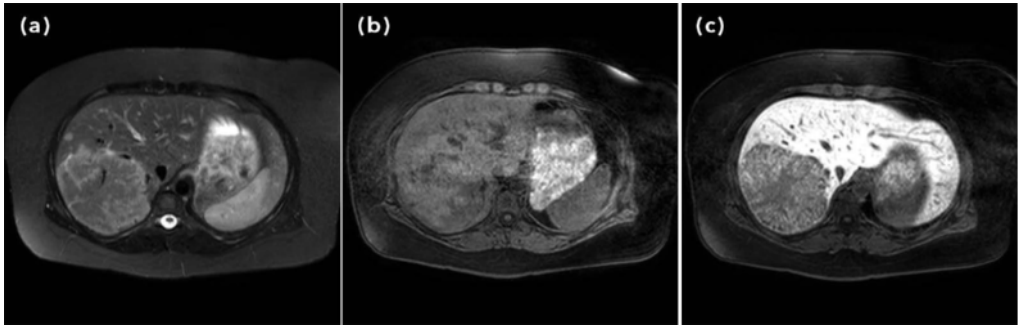
There was substantial agreement ( $\kappa = 0.64$ ) between the two readers concerning assessment of hypointensity versus iso- or hyperintensity in the hepatocyte contrast phase. In the hepatocyte phase, 63 % of the inflammatory HCAs (12/17 with gadobenate dimeglumine and 2/5 with gadoxetate disodium) showed areas of iso- or hyperintensity to the surrounding liver. From the 12 inflammatory HCAs scanned with gadobenate dimeglumine which were iso- to hyperintense to the liver in the hepatocyte phase, two had 1–25%, five had 76–99% and five had 100% iso- or hyperintensity to the liver (Figure 3). From the two inflammatory HCAs scanned with gadoxetate disodium which showed areas of iso- or hyperintensity to the liver in the hepatocyte phase, one had 1–25% and the other 26–50% iso- or hyperintensity (Figure 4).



**Figure 2.** A significant proportion of inflammatory HCAs are iso- or hyperintense to the liver in the hepatocyte phase after both gadobenate dimeglumine and gadoxetate disodium injection. In contrast to gadoxetate disodium, inflammatory HCAs scanned with gadobenate dimeglumine show larger surface enhancement. % = percentage of surface of HCA that is iso- or hyperintense in the hepatocyte phase; HCA = hepatocellular adenoma; dim. = dimeglumine; dis. = disodium; iso = iso- or hyperintense; hypo = hypointense.

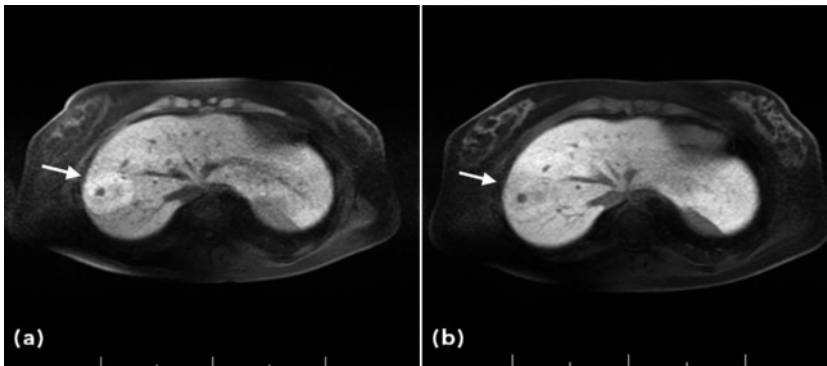


**Figure 3.** Isointensity of inflammatory HCA. (a) T2-weighted, fat suppressed axial image of a patient with 3 inflammatory HCAs (FATSAT FRFSE, 8 mm, 6315/93, flip angle = 90°); all three lesions show the bright rim that is typical of inflammatory HCAs. (b) T1-weighted precontrast axial image (GRE, 4 mm, 3.5/1.3, flip angle = 12°) on the same level, showing near isointensity to slight hyperintensity of the lesions. (c) One hour after injection of gadobenate dimeglumine, these lesions are barely visible due to isointensity compared to the surrounding liver.



**Figure 4.** A bright rim typical for inflammatory HCA. (a) T2-weighted axial image of the liver with a typical hyperintense rim (SPAIR, 5 mm, 2976/70, flip angle = 90°). Corresponding T1-weighted axial image in the precontrast phase (b) and 20 minutes after injection of gadoxetate disodium (c). Stellar contrast uptake is clearly visible.

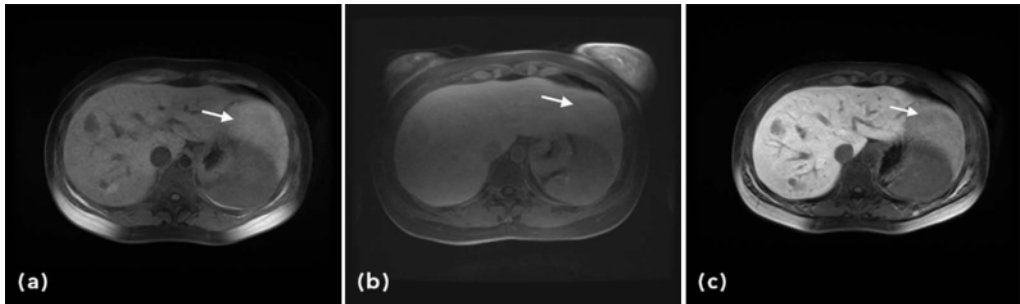
None of the non-inflammatory HCAs showed iso- or hyperintensity to the surrounding liver in the hepatocyte phase (five scanned with gadobenate dimeglumine and three with gadoxetate disodium). The inflammatory HCA that was scanned both with gadobenate dimeglumine and gadoxetate disodium showed isointensity to the liver in the hepatocyte phase after injection of gadobenate dimeglumine and showed hypointensity to the liver in the hepatocyte phase after injection of gadoxetate disodium.



**Figure 5.** Intrinsic hyperintensity. T1-weighted axial precontrast (a, arrow) and hepatocyte phase (b, arrow) scan (GRE, 5 mm, 2.7/1.2, flip angle = 12°) of an inflammatory HCA that is hyperintense relative to the surrounding liver before contrast and iso-intense in the hepatocyte phase. There was no uptake of contrast by absolute measurement; iso-intensity in the hepatocyte phase is due to intrinsic hyperintensity, as shown in the precontrast image. Gadobenate dimeglumine was used as the contrast agent.

We found that 10 of the inflammatory HCAs were hyperintense on the scan before contrast (eight scanned with gadobenate dimeglumine (Figure 5) and one with gadoxetate disodium and one with both (Figure 6)). This could be appreciated on the fatsaturated T1-weighted images but in four cases also on the in-phase T1-weighted sequence (three scanned with gadobenate dimeglumine and one with gadoxetate disodium).

All non-inflammatory HCAs were iso- to hypointense on all the T1-weighted sequences before contrast injection.



**Figure 6.** MRI of the liver of a patient who underwent both gadobenate dimeglumine and gadoxetate disodium injection. The lesion in segment 2 has an intrinsic hyperintensity on the scan before contrast (a, arrow). During the hepatocyte phase this lesion becomes isointense to the surrounding liver after injection of gadobenate dimeglumine (b, arrow) but clearly hypointense after injection of gadoxetate disodium (c, arrow). Since this lesion has another sign which typically corresponds with inflammatory HCA (atoll sign, not shown) it was stated as an inflammatory HCA, which was finally proved by biopsy.

## Discussion

In this preliminary investigation, we showed that the majority of inflammatory HCA can show areas of iso- or hyperintensity to the surrounding liver on hepatocyte phase imaging. A new molecular and pathological classification of HCA was recently established by the Bordeaux group [15]. This classification divides HCAs into four subgroups, according to genotypic/phenotypic characteristics and clinical features. The most prevalent subgroup consists of inflammatory HCAs [6;16;17] and this subtype, accounting for up to 50% of all HCAs, shows immunohistochemical expression of SAA and CRP. Inflammatory HCAs can exhibit telangiectatic features [18] and were previously categorized as a subtype of FNHs due to the typical presence of internal bile ducts [19]. However, as they show a molecular pattern closer to HCA, and may potentially be  $\beta$ -catenin positive – a suggested marker for potential malignant degeneration [12;19] – they are now recognised as HCAs.

To our knowledge, there are currently no studies describing the manner of presentation of these lesions on hepatocyte phase imaging. Grazioli et al. have reported that all included HCAs were hypointense to the liver in the hepatocyte phase after injection of gadobenate dimeglumine [11], notwithstanding that some of these lesions were probably inflammatory HCAs. This finding clearly contrasts with our findings. A possible explanation is that Grazioli et al. incorrectly identified inflammatory HCAs as FNHs by histology, as they harboured bile duct proliferations. The new classification of HCAs, as used in our study, is based on immunohistochemistry and has much greater accuracy and reproducibility, as earlier reported [13].

Various reasons could explain the specific iso- or hyperintense appearance of inflammatory HCAs in the hepatocyte phase after gadobenate dimeglumine. The most obvious being the late hepatocellular excretion in the bile ducts in inflammatory HCAs. However, a simple enhancement curve of gadobenate dimeglumine could not be applied because the patient had to return to the MRI scanner to be rescanned after 1 hour. Even so, subtraction techniques could not be used since signal intensities of the tissues may differ once a patient is placed back in the scanner. In addition to the interpretation of relative intensity compared to the surrounding liver, we used a formula to quantify the absolute uptake of contrast in the hepatocyte phase. We related the lesion signal intensities to those of the back musculature, in order to derive reliable comparable intensities. Evenso, this technique suffered from inconsistencies due to inherent B0 and B1 field inhomogeneities and we were not able to produce reliable enhancement curves. Thus, whether there was uptake in inflammatory HCAs could not be correctly analyzed.

Another finding can probably also explain the iso- or hyperintensity on hepatocyte phase imaging, and that is the hyperintensity (intrinsically or due to steatosis) before contrast of most of the inflammatory HCAs (Figure 6) in contrary to the non-inflammatory HCAs. Since the hepatocyte phase uptake of gadobenate dimeglumine is only 5% in the liver, the lesions could still be hyperintense when they harbour intrinsic hyperintensity (on the in-phase sequence and the fat saturated sequence) or are surrounded by steatosis (on the fat saturated sequence).

A third reason could also explain our findings in the gadobenate dimeglumine group. As mentioned by Grazioli et al, FNHs and HCAs are more easily discernable from each other after 3 hours, in comparison to 1 hour post-injection [11]. In our preliminary report, all patients were scanned 1 hour postinjection. Possibly waiting up to 3 hours all inflammatory HCAs would become hypointense to the liver. However, we did not perform a scan delay of more than 1 hour, and thus this argument can only be speculative.

Two of the inflammatory HCAs, and non of the non-inflammatory HCAs were iso- to hyperintense relatively to the liver 20 minutes after injection of gadoxetate disodium. This can be conferred to absolute enhancement since the patients were not replaced during the examination time. Recently, Yoneda et al also mentioned a case of a  $\beta$ -catenin-positive HCA which showed

enhancement on the hepatobiliary phase [20]. This was found to be the result of intense OATP 8 expression, which was verified to be the main uptake transporter of gadoxetate disodium into HCC cells and hepatocytes [21]. Our cases did not show  $\beta$ -catenin expression but were positive for SAA and CRP expression making it typical inflammatory HCAs. It has however to be stressed that the expression of  $\beta$ -catenin can be only focal making it prone to false negative findings. Furthermore, inflammatory HCAs may contain  $\beta$ -catenin expression as we showed earlier in a previous study [13]. We did not analyze OATP 8 expression in our two cases.

Since the percentage of uptake of contrast in hepatocytes is significantly different between the two contrast agents, we have the impression that when using gadoxetate disodium inflammatory HCAs become more hypointense to the surrounding liver in the hepatocyte phase. This is our impression in different clinical cases, but also proven in one patient who underwent both contrast injections. Probably, as was shown in this case, the T1 intrinsic hyperintensity of some inflammatory HCAs is not enough suppressed in the hepatocyte phase when using gadobenate dimeglumine. However, further studies where both contrast agents are used are warranted in order to draw conclusions on this issue.

Our findings may have important implications: since inflammatory HCAs are the most prevalent subtype of HCA and may be prone to degeneration and bleeding, all imaging characteristics should be thoroughly analyzed and all elements of the scan, including T1-weighted, T2-weighted and dynamic contrast characteristics, should be taken into account. While inflammatory HCAs typically present with an atoll sign [6], FNHs tend to show more intensive enhancement in comparison to HCAs [5]. A 'spoke wheel' scar configuration is typically seen in FNHs [5]. In addition, T1-hyperintensity due to blood degradation products is a typical sign of a HCA and is seldom seen in FNH [1]. Even if a lesion is hyperintense or isointense on both gadobenate dimeglumine or gadoxetate disodium, other imaging characteristics (as mentioned above) have to be carefully examined to allow clear differentiation between FNH and inflammatory HCA.

The proportion of inflammatory HCA was quite high in our study population. However, recent insights showed that these lesions are probably the most prevalent subgroup accounting for more than half of the HCAs [13].

The major limitations of this study were the retrospective design and the relatively low number of patients included. Although our results represent a valuable first step, a larger patient group is required to more accurately analyze the percentage of inflammatory HCAs that shows areas of iso- or hyperintensity to the surrounding liver.

In conclusion, inflammatory HCA can show areas of iso- or hyperintensity on hepatocyte phase when using gadobenate dimeglumine in MR imaging, and therefore other typical imaging features should also be used when differentiating between HCA and FNH.



## References

1. Hussain SM, van den Bos IC, Dwarkasing RS, Kuiper JW, den Hollander J. Hepatocellular adenoma: findings at state-of-the-art magnetic resonance imaging, ultrasound, computed tomography and pathologic analysis. *Eur Radiol* 2006;16(9):1873-1886.
2. Nagorney DM. Benign hepatic tumors: focal nodular hyperplasia and hepatocellular adenoma. *World J Surg* 1995;19(1):13-18.
3. van Aalten SM, Terkivatan T, de Man RA, et al. Diagnosis and treatment of hepatocellular adenoma in the Netherlands: similarities and differences. *Dig Surg* 2010;27(1):61-67.
4. Ronot M, Bahrami S, Calderaro J, et al. Hepatocellular adenomas: accuracy of magnetic resonance imaging and liver biopsy in subtype classification. *Hepatology* 2011;53(4):1182-1191.
5. Hussain SM, Terkivatan T, Zondervan PE, et al. Focal nodular hyperplasia: findings at state-of-the-art MR imaging, US, CT, and pathologic analysis. *Radiographics* 2004;24(1):3-17; discussion 18-19.
6. van Aalten SM, Thomeer MG, Terkivatan T, et al. Hepatocellular adenomas: correlation of MR imaging findings with pathologic subtype classification. *Radiology* 2011;261(1):172-181.
7. Fidler J, Hough D. Hepatocyte-specific magnetic resonance imaging contrast agents. *Hepatology* 2011;53(2):678-682.
8. Goodwin MD, Dobson JE, Sirlin CB, Lim BG, Stella DL. Diagnostic challenges and pitfalls in MR imaging with hepatocyte-specific contrast agents. *Radiographics* 2011;31(6):1547-1568.
9. Nakamura Y, Tashiro H, Nambu J, et al. Detectability of hepatocellular carcinoma by gadoxetate disodium-enhanced hepatic MRI: tumor-by-tumor analysis in explant livers. *J Magn Reson Imaging* 2013;37(3):684-691.
10. van Kessel CS, de Boer E, Kate FJ, Brosens LA, Veldhuis WB, van Leeuwen MS. Focal nodular hyperplasia: hepatobiliary enhancement patterns on gadoxetic-acid contrast-enhanced MRI. *Abdom Imaging* 2012.
11. Grazioli L, Morana G, Kirchin MA, Schneider G. Accurate differentiation of focal nodular hyperplasia from hepatic adenoma at gadobenate dimeglumine-enhanced MR imaging: prospective study. *Radiology* 2005;236(1):166-177.
12. Bioulac-Sage P, Rebouissou S, Sa Cunha A, et al. Clinical, morphologic, and molecular features defining so-called telangiectatic focal nodular hyperplasias of the liver. *Gastroenterology* 2005;128(5):1211-1218.
13. van Aalten SM, Verheij J, Terkivatan T, Dwarkasing RS, de Man RA, IJzermans JN. Validation of a liver adenoma classification system in a tertiary referral centre: implications for clinical practice. *J Hepatol* 2011;55(1):120-125.
14. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33(1):159-174.
15. Bioulac-Sage P, Laumonier H, Couchy G, et al. Hepatocellular adenoma management and phenotypic classification: the Bordeaux experience. *Hepatology* 2009;50(2):481-489.
16. Laumonier H, Bioulac-Sage P, Laurent C, Zucman-Rossi J, Balabaud C, Trillaud H. Hepatocellular adenomas: magnetic resonance imaging features as a function of molecular pathological classification. *Hepatology* 2008;48(3):808-818.
17. Paradis V, Champault A, Ronot M, et al. Telangiectatic adenoma: an entity associated with increased body mass index and inflammation. *Hepatology* 2007;46(1):140-146.
18. Bioulac-Sage P, Rebouissou S, Thomas C, et al. Hepatocellular adenoma subtype classification using molecular markers and immunohistochemistry. *Hepatology* 2007;46(3):740-748.
19. Paradis V, Benzekri A, Dargere D, et al. Telangiectatic focal nodular hyperplasia: a variant of hepatocellular adenoma. *Gastroenterology* 2004;126(5):1323-1329.
20. Yoneda N, Matsui O, Kitao A, et al. Beta-catenin-activated hepatocellular adenoma showing hyperintensity on hepatobiliary-phase gadoxetic-enhanced magnetic resonance imaging and overexpression of OATP8. *Jpn J Radiol* 2012;30(9):777-782.



21. Kitao A, Zen Y, Matsui O, et al. Hepatocellular carcinoma: signal intensity at gadoxetic acid-enhanced MR Imaging--correlation with molecular transporters and histopathologic features. *Radiology* 2010;256(3):817-826.



# Chapter 7

## Genotype-phenotype correlations in hepatocellular adenoma: an update of MRI finding

Maarten G. Thomeer

Mirelle E. E. Bröker

Quido de Lussanet

Katharina Biermann

Roy S. Dwarkasing

Rob de Man

Jan N. IJzermans

Marianne de Vries

## Abstract

Hepatocellular adenoma (HCA) is a generally benign liver entity with a potential for malignancy and bleeding. HCA are categorized into four subtypes on the basis of the genetic and pathologic features: (a) hepatocyte nuclear factor 1  $\alpha$ -mutated HCA, (b)  $\beta$ -catenin-mutated HCA, (c) inflammatory HCA, and (d) unclassified HCA. MRI plays an important role in the diagnosis, subtype characterization and detection of complications of HCA, and differentiation from focal nodular hyperplasia (FNH).

In this review, we present an overview of genetic abnormalities, oncogenesis, and typical and atypical MRI findings of specific subtypes of HCA on contrast enhanced MRI with or without hepatobiliary contrast agents (Gd-BOPTA-gadobenate dimeglumine and Gd-EOB-DTPA-gadoxetate disodium). We will discuss their different management implications after diagnosis.

# Introduction

Hepatocellular adenoma (HCA) is a rare, benign tumor of the liver that occurs predominantly in young and middle-aged women [1]. In contrast to focal nodular hyperplasia (FNH), HCA may involve complications such as a life-threatening bleeding and malignant degeneration [1-3]. The strong association between the occurrence of HCA and the use of oral contraceptives was first acknowledged in 1970's [4] and the incidence of HCA is now thought to be 30 times greater in oral contraceptive users as compared to non-users [5;6]. A dose-dependent association and spontaneous regression following withdrawal of oestrogens have also been described [4;7]. However, the exact role of oestrogen in HCA is still poorly understood.

In this review, we present an overview of typical and atypical MRI findings of different HCAs and with respect to FNH, and discuss various pitfalls that may be encountered on MRI.

## The new classification of HCAs

A molecular and immunohistochemical classification (IHC) of HCA has been introduced by the Bordeaux group [8;9] (Table 1) in which HCAs are divided into four subgroups based on clear genetic differences.

**Table 1.** Immunohistochemical and MRI signs used for differentiating different HCA subtypes. Note that in exceptional cases, GS can be normal in  $\beta$ -catenin HCAs. Also note that inflammatory HCAs may show  $\beta$ -catenin positivity. The MRI signs are preliminary and based on three recent papers.

Subtype HCA	LFABP - HCA	$\beta$ -catenin HCA	Inflammatory HCA	Unclassified HCA
Immunohistochemical staining				
GS	-	+/-	-	-
$\beta$ -catenin	-	+	+	-
CRP	-	-	+	-
SAA	-	-	+	-
LFABP	-	+	+	+
Typical MRI findings				
	- Diffuse homogenous lesional steatosis	- Faint scar	- Atoll sign - Strong diffuse hyperintense signal on T2-weighting	

HCA = hepatocellular adenoma; GS = glutamate synthetase; CRP = C-reactive protein; SAA = serum Amyloid A; LFABP = liver fatty acid binding protein.

The first group, accounting for 30–40 % of cases, is defined by the presence of Hepatocyte Nuclear Factor 1 $\alpha$  (HNF1A) mutations [10]. The HNF1A gene controls lipid metabolism and mediates the downregulation of liver fatty acid binding protein (LFABP). LFABP downregulation is typically observed using LFABP staining, which has an accuracy of 100% [8;11]. The most

typical presentation of group 1 (i.e. LFABP-) HCA lesions is the aberrant presence of internal steatosis. It should be noted, however, that internal steatosis is not sufficient for diagnosing this subtype, since other subtypes may exhibit internal steatosis as well [12]. We prefer to avoid the term steatotic HCA, used in some literature for this particular subtype [13]. Some of the patients with hepatocyte nuclear factor 1  $\alpha$ -mutated HCA have an associated mutation that is thought to be responsible for maturity-onset non-insulin dependent diabetes (MODY3) [14]. therefore , once the diagnosis of this subtype of HCA is made , clinicians should be warranted of the possibility of underlying diabetes.

A second group, 10–15% of the cases, shows the presence of activating mutations of  $\beta$ -catenin [15]. Whereas  $\beta$ -catenin is phosphorylated and degraded by proteasomes under physiological circumstances, tumors fail to downregulate  $\beta$ -catenin and instead show nuclear accumulation of the protein [16]. This accumulation is known to trigger an important signalling pathway in several cancers, but is typified by hepatocellular carcinoma [17]. Although  $\beta$ -catenin activation is too insensitive for IHC due to its visibility in only a few sporadic nuclei [12], another product of the same  $\beta$ -catenin activation pathway, glutamine synthetase (GS), is homogeneously expressed in most lesions and can therefore be used as a sensitive and specific immunohistochemical staining for  $\beta$ -catenin HCAs [8;12]. It should be noted that focal nodular hyperplasia (FNH) also presents with a diffuse GS staining. Differentiation of these two benign lesions ( $\beta$ -catenin HCA and FNH) is not difficult on large resections, as the map-like staining in FNH contrasts with the homogeneous expression in  $\beta$ -catenin HCAs [12]. Nonetheless, it can be challenging for the pathologist to make this differentiation in small biopsies.

The third group, 40–50% of all HCAs, shows inflammatory expression [18]. A typical feature of these lesions is the activation of acute phase inflammation proteins such as serum amyloid A and C-reactive protein [8]. This subgroup is also linked to obesity and high alcohol intake [19]. Importantly, these lesions are associated with homogenous GS and  $\beta$ -catenin staining and therefore it has been described they could have an increased risk of malignant degeneration [12].

The final group, accounting for 10–25% of the HCAs, shows no specific genetic alterations and are therefore currently referred to as ‘unclassified’ [16].

### Implication for diagnosis

While IHC (of HNFA,  $\beta$ -catenin, GS, SAA and CRP) has proven to be very effective in differentiating the four types of HCA, it is also useful in the differentiation between HCA and FNH [9;20]. In a retrospective, multicentre study in France, Bioulac-Sage et al found that the certainty of biopsy diagnosis of FNH increased from 53% to 87% when additional IH and GS staining was used [20]. The certainty of biopsy diagnosis of HCA also increased from 59% to 74% when IH analyses were used. Prior to the introduction of these markers HCA was oftentimes

misdiagnosed as a FNH during histological examination, especially when inflammatory HCA (formerly known as teleangiectatic FNH) were involved [19]. These latter lesions include the bile duct proliferation seen in FNH, but show the behaviour of an HCA (including the described risk of malignant degeneration and bleeding). This confusion should be taken into account when evaluating older radiologic descriptions of HCA and FNH, as the reference standard has only become significantly more accurate since the introduction of these markers.

Conventional MRI findings for HCAs

HCAs are primarily diagnosed by non-invasive imaging techniques [1], and typical MRI characteristics can be used for differentiating HCA from FNH. According to recent literature, some MRI findings are more typical than others (Table 2). T1-weighted hyperintensity seems to be prevalent only in HCAs and is probably caused by blood degeneration products or glycogen storage [21-23].

**Table 2.** Classical signs that can be used for differentiating HCAs from FNHs. We distinguish strong signs from weak signs. Strong signs are defined to be characteristic for that lesion. Weak signs are more common in either HCA or FNH, but can occur in both.

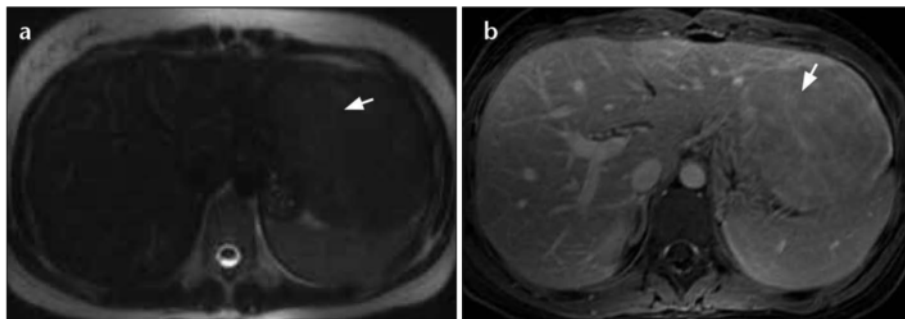
HCA	FNH
Strong signs	
Strong hyperintensity on T2 -weighting	Spoke wheel appearance of scar
Hyperintensity on T1-weighting	
Cystic parts	
Haemorrhagic parts	
Diffuse intralesional steatosis	
Atoll sign	
Weak signs	
Faint arterial enhancement	Scar
Liver steatosis	Lobular contours
Multiple lesions	Strong arterial enhancement

HCA = hepatocellular adenoma; FNH = focal nodular hyperplasia.

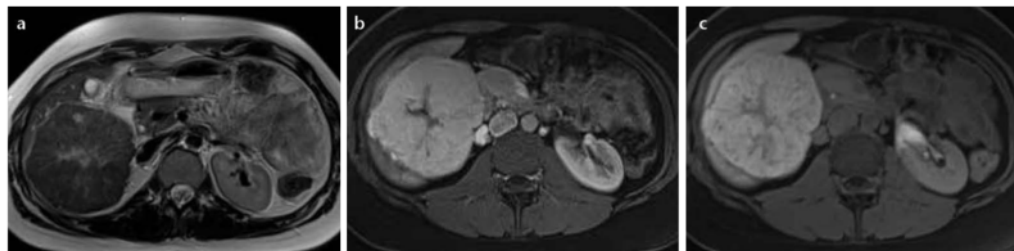
Other signs of non-neoplastic degeneration can be appreciated in some HCAs through the visualization of internal bleeding cysts, necrosis or fluid. To the best of our knowledge, these finding have not been described for FNHs [2]. A strong T2-weighted hyperintense band in peripheral areas of the lesion is a typical sign for HCAs, and was only (possibly) visible in one FNH in our series to date (see below) [21].

The finding of a central scar is a more commonly described imaging characteristic of FNH [2]. However, we also found linear central scars in 21% of proven HCAs [21]. This sign, characterized by a T2-weighted central scar with late enhancement on delayed phase, does not seem to be sufficiently robust to allow differentiation of these two lesions. In addition,  $\beta$ -catenin-associated HCAs appear to have a faint central scar in up to 75 % of cases [21] (Figure 1). On the other

hand, in our experience a typical 'spoke wheel' appearance of a central scar is in fact only visible in FNH (Figure 2) [2].



**Figure 1.** Axial T2-weighted (a) and T1-weighted (b) MR images of the liver after contrast injection. Histologically,  $\beta$ -catenin staining is positive in this patient. This faint scarlike region is T2 hyperintense (a arrow) with late enhancement after contrast injection of a nonspecific Gadolinium based contrast agent (b arrow), findings similar to those expected in FNHs. In our opinion, a lesion with a scar but lacking a spoke wheel aspect is not only typical of FNH, but may also occur in HCAs and more typically, in  $\beta$ -catenin HCAs. HCA = hepatocellular adenoma; FNH = focal nodular hyperplasia.



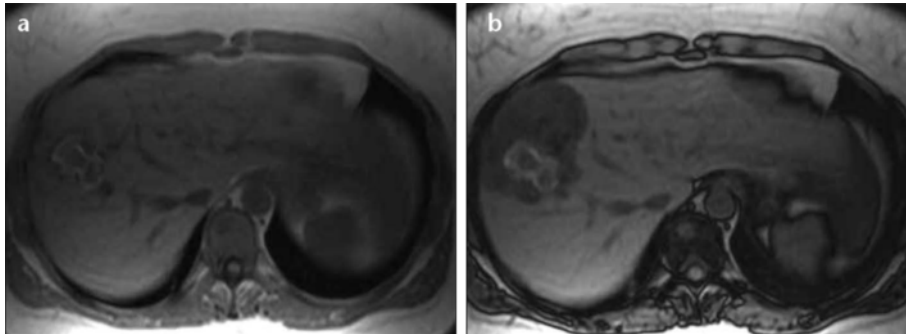
**Figure 2.** Axial T2-weighted (a), venous phase T1-weighted and 20 minutes T1-weighted MR images of a patient with typical 'spoke wheel' aspect typical for a FNH. A central scar with divergences to the periphery is visible on T2-weighting, reminiscent of a spoke wheel (a). These so-called 'spokes' are enhancing in the venous phase normally using an aspecific contrast agent or gadobenate dimeglumine. But when using gadoxetate disodium, the central scar and the spokes stay hypointense due to pseudo wash-out (b). After 20 minutes the majority of the lesion (except the scar) becomes hyperintense owing internal bile duct proliferation (c). FNH = focal nodular hyperplasia.

The surrounding liver steatosis, intralesional fat accumulation, faint arterial enhancement pattern and the finding of multiple lesions are more or less typical for HCA, but can also occur in FNHs [21-23]. A lobular border is oftentimes described as a typical sign for the diagnosis FNH, but we and others have found that the occurrence of a lobular border is not uncommon in HCAs [21;22].



### MRI findings based on the new subclassification

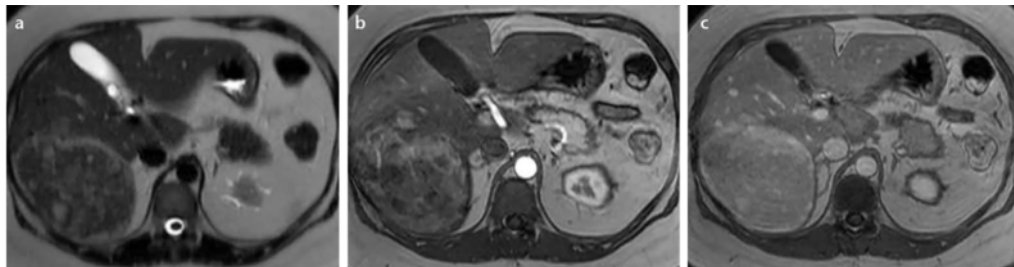
Laumonier et al were the first to publish on the typical MRI features of HCA according to the subgroup classification [23]. A homogeneous dropout of signal on the T1-weighted out-of-phase sequence had a sensitivity of 86.7% and a specificity of 100% for LAFBP – HCAs (Figure 3), while this drop-out was absent or only focal (heterogeneous) in inflammatory HCAs. Moreover, marked hyperintensity on T2-weighted sequences was found to be typical for inflammatory HCAs, with a sensitivity of 85.2% and a specificity of 87.5%.



**Figure 3.** Axial T1-weighted in (a) and opposed (b) phase MR image of a LAFBP – HCA. This lesion shows a typical diffuse and homogeneous suppression of signal in the lesion due to fat accumulation. Notice also the bleeding residue centrally in the lesion as T1-weighted hyperintense zone. Besides steatosis, this lesion showed LAFBP staining on histology – a very sensitive marker for HNF1A mutation. HCA: hepatocellular adenoma, LAFBP = liver fatty acid binding protein; HNF1A = hepatocyte nuclear factor 1A.

We found similar presentations for LAFBP – HCAs and inflammatory HCAs [21]. In addition, we showed that a hyperintense rim on T2-weighted sequences was diagnostic for inflammatory HCAs. This hyperintense rim sign corresponds to sinusoidal dilatation (Figure 4) and is also referred to as ‘atoll sign’. This includes a hyperintense rim in the periphery of the lesion on T2-weighted imaging, with an isointensity in the centre of the lesion reminiscent of the sea within an atoll [24]. Small intralesional T2-hyperintense nodules can be found in the centre of the lesion (small islands) [21].

Several authors have reported that a faint scar may be a possible sign in  $\beta$ -catenin HCA [21;23;25], but the number of published  $\beta$ -catenin HCAs to date is too low to draw any firm conclusions.



**Figure 4.** Axial T2-weighted (a), arterial (b) and venous (c) T1-weighted MR images of the liver in a patient with an inflammatory HCA. Typical presentation of an inflammatory HCA. The lesion presents with an atoll sign, which appears as a T2-weighted hyperintense rim (peripheral island with central sea) with or without central hyperintense islands as can be found inside an atoll (a). The lesion is hypervascular in the arterial phase (b), with late enhancement of the peripheral rim and central islands (c). On histology, these lesions have a positive immunostaining with inflammatory parameters including CRP and SAA. The T2-weighted hyperintense rim is thought to be caused by local peliosis. HCA = hepatocellular adenoma; CRP = C-reactive protein; SAA = serum amyloid A

### MRI findings using liver-specific contrast agents

With the introduction of hepatobiliary contrast agents, an important tool became available for differentiating HCAs from FNH [26].

Two gadolinium-based contrast agents are currently available: gadobenate dimeglumine (Multihance, Milan, Italy and gadoxetate disodium (Primovist, Berlin, Germany)(US brand name, Eovist).

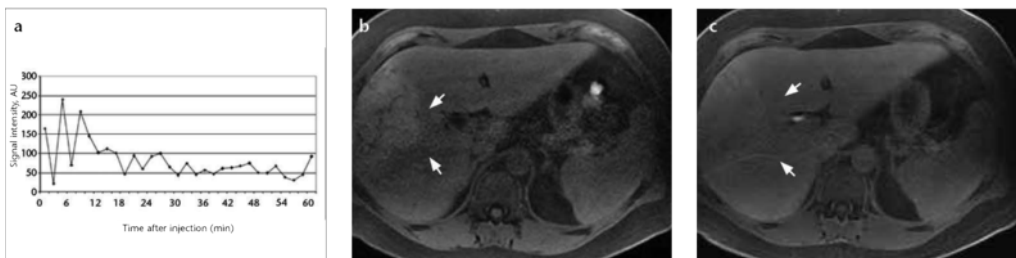
Both agents show hepatocyte uptake and biliary excretion, with a hyperintense liver in the hepatocyte phase on T1-weighted imaging as a consequence. Lesions which contain bile ducts also seem to enhance. This is typically the case in FNHs and in some hepatocellular carcinomas (HCCs) [26]. The most important finding is that HCAs do not normally show hepatocyte uptake and biliary excretion and are therefore hypointense to the liver in the hepatocyte phase.

There are strong arguments for uptake of gadoxetate disodium in FNHs and some HCCs by organic anion transporter polypeptide (OATP) channels [25;27;28]. However one study suggest other mechanism of uptake in the liver of both gadobenate dimeglumine and gadoxetate disodium [29].

Yoneda et al found that OATP8 is present in the periphery and not in the center of FNH, explaining the peculiar aspect of some FNH in the hepatocyte phase (ring-enhancement-type) [27].

Gadobenate dimeglumine is excreted by the liver at a significantly lower percentage (5%) than gadoxetate disodium (50%) [26], with the result that gadoxetate disodium produces a significantly greater signal intensity change in the hepatocyte phase than gadobenate dimeglumine. This greater effect can be helpful in cases where liver activity, and thus changes in enhancement are low, as seen in cirrhosis. A cause for this lower liver enhancement is the competitive uptake of these contrast agents in relation to bilirubin, the latter being enhanced in patients with liver cirrhosis. In addition, lesions that have an intrinsic T1-weighted hyperintensity, as frequently seen in HCAs, may often show insufficient liver enhancement when using gadobenate dimeglumine, resulting in the hyperintense HCA becoming isointense instead of hypointense to the surrounding liver [30].

Whether there is a difference in uptake of hepatobiliary contrast agents between the different subtypes of HCA is still largely unknown. We recently reported some initial results in which we noted that inflammatory HCAs range from isointense to hyperintense in relation to the liver in the hepatocyte phase (Figure 5) [30]. However, the underlying cause of this finding was not obvious. Of particular interest is the fact that, in contrast to other HCAs, inflammatory HCAs often harbour internal bile ducts, possibly explaining the late isointensity in the hepatocyte phase. However, this phenomenon is difficult to discern from isointensity due to intrinsic hyperintensity prior to contrast



**Figure 5.** Enhancement curve (a) of a proven inflammatory HCA who was scanned every 2.5 minutes over a period of one hour following the injection of gadobenate dimeglumine (a). This approach allows an enhancement curve to be reproduced that shows no late enhancement. The lesion remains otherwise isointense after 1 hour on the axial fat-saturated T1-weighted MR image (c arrow), due to intrinsic hyperintensity, as can be appreciated on the axial fat-saturated T1-weighted sequence prior to contrast injection (b). HCA = hepatocellular adenoma

### Implications for patient care

Although the use of either gadobenate dimeglumine or gadoxetate disodium may often be indicated, it can be difficult for a radiologist to choose between these two contrast agents. The major differences are listed in Table 3.

**Table 3.** Comparison of the hepatobiliary contrast agents gadobenate dimeglumine and gadoxetate disodium for liver imaging.

	Gadobenate dimeglumine	Gadoxetate disodium
T1-effect	+++	+
Cost	+	++
Shortness of investigation	+	++(+)
Differentiation of HCA from FNH	+	+
Differentiation of benign liver lesions (haemangiomas)	++	-(+)

HCA = hepatocellular adenoma; FNH = focal nodular hyperplasia.

As already noted by others, gadoxetate disodium might be preferred over gadobenate dimeglumine when considering the rapid hepatobiliary phase at 20 minutes or earlier, whereas in patients scanned with gadobenate dimeglumine the hepatobiliary phase is obtained with a second MR possibly minimally one hour after the initial MRI and injection of contrast agent [24]. However, it is important to note that the early hepatobiliary excretion of gadoxetate disodium has the consequence that the dynamic phase cannot be separated from the hepatobiliary phase. This means that the interpretation of the dynamic phase in case of suspected haemangiomas is disturbing. Additionally, when using gadoxetate disodium, a large number of liver lesions become hypointense in the late dynamic phase due to pseudo wash-out rather than actual wash-out (due to the enhancement of the surrounding liver). Finally, radiologists should be aware that because of this pseudo wash-out the central scar in FNH may become hypointense instead of hyperintense in the late venous phase when using gadoxetate disodium.

Another argument in favour of gadobenate dimeglumine over gadoxetate disodium is the considerably lower cost of the former. This is probably due to the wider range of possible indications for using gadobenate dimeglumine, which appears to be a useful nonspecific gadolinium-based contrast agent. With the T1-effect of gadoxetate disodium being less pronounced, this agent is indicated for specific liver imaging only.

In case of an atypical liver lesion, where the differential diagnosis is broad including hemangioma and malignant tumours, a nonspecific contrast agent or gadobenate dimeglumine is generally preferred over gadoxetate disodium. In cases where MRI is prescribed solely for the differentiation of FNH and HCA, gadoxetate disodium should be sufficient. This is often the case when preliminary external data on CT or MRI are present.

Although contrast-enhanced ultrasound is beyond the scope of this review, it may be speculated that in the near future, differentiation between the different subtypes of HCAs may also become possible with the use of contrast sonography [31].

In addition, other markers may come to play a more prominent role in the differentiation of HCAs from FNHs or HCCs [32;33].

### **Malignant degeneration**

Malignant degeneration of HCA has been reported but seems to occur very rarely. In a recent meta-analysis, a total of 1635 HCAs were retrieved, reporting an overall frequency of malignant transformation in 4.2%. Most of these lesions were larger than 5 cm in diameter. However, the described overall frequency is probably an overestimation due to the limited sample sizes of the studies and as most studies only describe resected HCA [29].

Furthermore, it is interesting that, so called, malignant degenerated HCAs show a pattern of a nodule within a nodule, or two tumours lying adjacent to each other [28;30]. In this situation it might be questioned whether these nodules have been exposed to the same cellular changes.

The malignant potential of HCA seems to prevail more in B-catenin HCAs which are also more prevalent in men [15]. Bioulac-Sage et al. reported a prevalence of six HCCs of 128 proven HCAs, all of them being  $\beta$ -catenin HCAs whether inflammatory or not [15].

Management of these  $\beta$ -catenin mutated HCAs varies between hospitals from conservative to surgical management. Additional experiences from more hospitals are needed to correlate the classification system with clinical management [28].

### **Internal bleeding in HCA**

Rupture and bleeding have both been described in HCA, perhaps it is the hypervascular nature of these lesions that makes them more prone to bleeding. In contrast, FNHs do not show rupture or bleeding despite their hypervascularity. We found a prevalence for ruptured HCA of 16% in a meta-analysis [3], although the papers studied did show some evidence of selection bias as patients with silent HCAs were underreported due to the lack of any indication for imaging in most of these patients. The meta-analysis showed that larger lesions (larger than 5 cm) are more often involved. There also seems to be no difference in the prevalence of internal bleeding depending on the subtype of HCA. Analysing their own database, Bioulac-Sage et al. also found no difference in the chance of macroscopic bleeding between LFABP- HCAs and inflammatory HCAs, the two most prevalent subtypes (Figure 4) [15].

Since the actual risk of bleeding is reported to be higher in lesions larger than 5 cm, resection of these lesions may be warranted. Decisions related to the treatment of smaller HCAs are still complex and other characteristics such as male gender,  $\beta$ -catenin positivity, or the wish to become pregnant, can also play an important role. Larger studies will need to be performed to evaluate whether there is a correlation between subtypes and bleeding ruptures. Pregnant

patients with HCAs deserve special attention, as maternal and foetal mortality are reported to be not negligible [34;35;36]. We advocate patients with HCAs larger than 5 cm to be treated before they become pregnant. In cases where a HCA larger than 5 cm is found during pregnancy, data are currently too sparse to allow firm conclusions to be drawn, and we believe that current management in these patients should be individualized [36;37].

## **HCA in men**

HCA mostly occur in women in their second and third decade in life [5], and the incidence in men is very low. Most described cases of HCA in men occurred after the chronic intake of exogenous hormones [38] or in men with glycogen storage disease [39]. Because of its low incidence, diagnosis should be questioned and we advise to be more liberal performing a biopsy when these tumours are found in men, to exclude a pre-malignant or malignant liver tumor.

## **Conclusion**

Advances made these recent years greatly improved our understanding of HCAs, leading to the recognition of subtypes and better differentiation from FNH. Reports and conclusions drawn from MRI data in older studies, prior to the introduction of GS staining in particular, should therefore be evaluated with caution.

Gained experience in the atypical findings and potential pitfalls that may be encountered on MR imaging has added in the diagnosis of HCA. The introduction of hepatobiliary contrast agents has contributed even more, particularly in differentiating between HCA and FNA, which have now become valuable tools in daily clinical practice. However, the radiologist must always be alert for possible errors in diagnosis, owing to differences between hepatobiliary and nonspecific gadolinium based contrast agents. With respect to the treatment of HCA a number of recommendations can be made, however, other questions require further studies.

## References

1. Hussain SM, van den Bos IC, Dwarkasing RS, Kuiper JW, den Hollander J. Hepatocellular adenoma: findings at state-of-the-art magnetic resonance imaging, ultrasound, computed tomography and pathologic analysis. *Eur Radiol* 2006; 16:1873-1886. [CrossRef]
2. Hussain SM, Terkivatan T, Zondervan PE, et al. Focal nodular hyperplasia: findings at state-of-the-art MR imaging, US, CT, and pathologic analysis. *Radiographics* 2004; 24:3-19. [CrossRef]
3. van Aalten SM, de Man RA, IJzermans, Terkivatan T. Systematic review of haemorrhage and rupture of hepatocellular adenomas. *Br J Surg* 2012; 99:911-916. [CrossRef]
4. Buhler H, Pirovino M, Akobiantz A, et al. Regression of liver cell adenoma. A follow-up study of three consecutive patients after discontinuation of oral contraceptive use. *Gastroenterology* 1982; 82:775-782.
5. Reddy KR, Schiff ER. Approach to a liver mass. *Semin Liver Dis* 1993; 13:423-435. [CrossRef]
6. Rooks JB, Ory HW, Ishak KG, et al. Epidemiology of hepatocellular adenoma. The role of oral contraceptive use. *JAMA* 1979; 242:644-648. [CrossRef]
7. Steinbrecher UP, Lisbona R, Huang SN, Mishkin S. Complete regression of hepatocellular adenoma after withdrawal of oral contraceptives. *Dig Dis Sci* 1981; 26:1045-1050. [CrossRef]
8. Bioulac-Sage P, Rebouissou S, Thomas C, et al. Hepatocellular adenoma subtype classification using molecular markers and immunohistochemistry. *Hepatology* 2007; 46:740-748. [CrossRef]
9. Bioulac-Sage P, Balabaud C, Bedossa P, et al. Pathological diagnosis of liver cell adenoma and focal nodular hyperplasia: Bordeaux update. *J Hepatol* 2007; 46:521-527. [CrossRef]
10. Bluteau O, Jeannot E, Bioulac-Sage P, et al. Bi-allelic inactivation of TCF1 in hepatic adenomas. *Nat Genet* 2002; 32:312-315. [CrossRef]
11. Akiyama TE, Ward JM, Gonzalez FJ. Regulation of the liver fatty acid-binding protein gene by hepatocyte nuclear factor 1alpha (HNF1alpha). Alterations in fatty acid homeostasis in HNF1alpha-deficient mice. *J Biol Chem* 2000; 275:27117-27122.
12. van Aalten SM, Verheij J, Terkivatan T, Dwarkasing RS, de Man RA, IJzermans JN. Validation of a liver adenoma classification system in a tertiary referral centre: implications for clinical practice. *J Hepatol* 2011; 55:120-125. [CrossRef]
13. Ronot M, Bahrami S, Calderaro J, et al. Hepatocellular adenomas: accuracy of magnetic resonance imaging and liver biopsy in subtype classification. *Hepatology* 2011; 53:1182-1191. [CrossRef]
14. Jeannot E, Mellottee L, Bioulac-Sage P, et al. Spectrum of HNF1A somatic mutations in hepatocellular adenoma differs from that in patients with MODY3 and suggests genotoxic damage. *Diabetes* 2010; 59:1836-1844. [CrossRef]
15. Bioulac-Sage P, Laumonier H, Couchy G, et al. Hepatocellular adenoma management and phenotypic classification: the Bordeaux experience. *Hepatology* 2009; 50:481-489. [CrossRef]
16. Nault JC, Bioulac-Sage P, Zucman-Rossi J. Hepatocellular benign tumors-from molecular classification to personalized clinical care. *Gastroenterology* 2013; 144:888-902. [CrossRef]
17. Cadoret A, Ovejero C, Saadi-Kheddouci S, et al. Hepatomegaly in transgenic mice expressing an oncogenic form of beta-catenin. *Cancer Res* 2001; 61:3245-3249.
18. Bioulac-Sage P, Blanc JF, Rebouissou S, Balabaud C, Zucman-Rossi J. Genotype phenotype classification of hepatocellular adenoma. *World J Gastroenterol* 2007; 13:2649-2654.
19. Bioulac-Sage P, Rebouissou S, Sa Cunha A, et al. Clinical, morphologic, and molecular features defining so-called telangiectatic focal nodular hyperplasias of the liver. *Gastroenterology* 2005; 128:1211-1218. [CrossRef]
20. Bioulac-Sage P, Cubel G, Taouji S, et al. Immunohistochemical markers on needle biopsies are helpful for the diagnosis of focal nodular hyperplasia and hepatocellular adenoma subtypes. *Am J Surg Pathol* 2012; 36:1691-1699. [CrossRef]
21. van Aalten SM, Thomeer MG, Terkivatan T, et al. Hepatocellular adenomas: correlation of MR imaging findings with pathologic subtype classification. *Radiology* 2011; 261:172-181. [CrossRef]

22. Bieze M, van den Esschert JW, Nio CY, et al. Diagnostic accuracy of MRI in differentiating hepatocellular adenoma from focal nodular hyperplasia: prospective study of the additional value of gadoxetate disodium. *AJR Am J Roentgenol* 2012; 199:26–34. [CrossRef]
23. Laumonier H, Bioulac-Sage P, Laurent C, Zucman-Rossi J, Balabaud C, Trillaud H. Hepatocellular adenomas: magnetic resonance imaging features as a function of molecular pathological classification. *Hepatology* 2008; 48:808–818. [CrossRef]
24. Grazioli L, Morana G, Kirchin MA, Schneider G. Accurate differentiation of focal nodular hyperplasia from hepatic adenoma at gadobenate dimeglumine-enhanced MR imaging: prospective study. *Radiology* 2005; 236:166–177. [CrossRef]
25. Yoneda N, Matsui O, Kitao A, et al. Beta-catenin-activated hepatocellular adenoma showing hyperintensity on hepatobiliary-phase gadoxetic-enhanced magnetic resonance imaging and overexpression of OATP8. *Jpn J Radiol* 2012; 30:777–782. [CrossRef]
26. Fidler J, Hough D. Hepatocyte-specific magnetic resonance imaging contrast agents. *Hepatology* 2011; 53:678–682. [CrossRef]
27. Yoneda N, Matsui O, Kitao A, et al. Hepatocyte transporter expression in FNH and FNH-like nodule: correlation with signal intensity on gadoxetic acid enhanced magnetic resonance images. *Jpn J Radiol* 2012; 30:499–508. [CrossRef]
28. Narita M, Hatano E, Arizono S, et al. Expression of OATP1B3 determines uptake of Gd-EOB-DTPA in hepatocellular carcinoma. *J Gastroenterol* 2009; 44:793–798. [CrossRef]
29. Pascolo L, Cupelli F, Anelli PL, et al. Molecular mechanisms for the hepatic uptake of magnetic resonance imaging contrast agents. *Biochem Biophys Res Commun* 1999; 257:746–752. [CrossRef]
30. Thomeer MG, Willemssen FE, Biermann KK, et al. MRI features of inflammatory hepatocellular adenomas on hepatocyte phase imaging with liver-specific contrast agents. *J Magn Reson Imaging* 2013 doi: 10.1002/jmri.24281. [CrossRef]
31. Laumonier H, Cailliez H, Balabaud C, et al. Role of contrast-enhanced sonography in differentiation of subtypes of hepatocellular adenoma: correlation with MRI findings. *AJR Am J Roentgenol* 2012; 199:341–348. [CrossRef]
32. Wang HL, Anatelli F, Zhai QJ, Adley B, Chuang ST, Yang XJ. Glypican-3 as a useful diagnostic marker that distinguishes hepatocellular carcinoma from benign hepatocellular mass lesions. *Arch Pathol Lab Med* 2008; 132:1723–1728.
33. Palmer PE, Christopherson WM, Wolfe HJ. Alpha1-antitrypsin, protein marker in oral contraceptive-associated hepatic tumors. *Am J Clin Pathol* 1977; 68:736–739.
34. van der Windt DJ, Kok NF, Hussain SM, et al. Case-orientated approach to the management of hepatocellular adenoma. *Br J Surg* 2006; 93:1495–1502. [CrossRef]
35. Noels JE, van Aalten SM, van der Windt DJ, et al. Management of hepatocellular adenoma during pregnancy. *J Hepatol* 2011; 54:553–558. [CrossRef]
36. van Aalten SM, Broker ME, Busschbach JJ, et al. Pregnancy and liver adenoma management: PALM-study. *BMC Gastroenterol* 2012; 12:82. [CrossRef]
37. Broker ME, IJzermans JN, van Aalten SM, de Man RA, Terkivatan T. The management of pregnancy in women with hepatocellular adenoma: a plea for an individualized approach. *Int J Hepatol* 2012; 2012:725–735. [CrossRef]
38. Psatha EA, Semelka RC, Armao D, Woosley JT, Firat Z, Schneider G. Hepatocellular adenomas in men: MRI findings in four patients. *J Magn Reson Imaging* 2005; 22:258–264. [CrossRef]
39. Howell RR, Stevenson RE, Ben-Menachem Y, Philyly RL, Berry DH. Hepatic adenomata with type 1 glycogen storage disease. *JAMA* 1976; 236:1481–1484. [CrossRef]



# Chapter 8

## **Letter to the editor: Quantitative analysis of hepatocellular adenoma and focal nodular hyperplasia in the hepatobiliary phase: external validation of LLCER method using gadobenate dimeglumine as contrast agent**

Maarten G Thomeer

Bibiche Gest

Hermen van Beek

Marianne De Vries

Roy Dwarkasing

Julia Klompenhouwer

Robert A De Man

Jan N IJzermans

Loes Braun

## To the editor:

We read the article by Roux et al. with interest [1]. Their purpose was to find a quantitative method to differentiate hepatocellular adenoma (HCA) from focal nodular hyperplasia (FNH), since visual comparison between liver intensity and lesion intensity in the hepatobiliary phase after gadobenate dimeglumine (Multihance, Bracco Imaging, Milan, Italy) is often inadequate [2]. Using a lesion-to-liver contrast enhancement ratio (LLCER) for differentiation between the two lesions, they found an optimal cut-off of -0.3 % for observer 1 and 1.2 % for observer 2. We performed an external validation of these findings.

In a retrospective analysis of 69 patients with 77 pathologically-proven liver lesions (49 HCA, 28 FNH), fat-suppressed T1-weighted sequences were acquired before and after (at least one hour) administering Gd-BOPTA contrast in the hepatobiliary phase. This study was approved by our institutional review board and informed consent was waived.

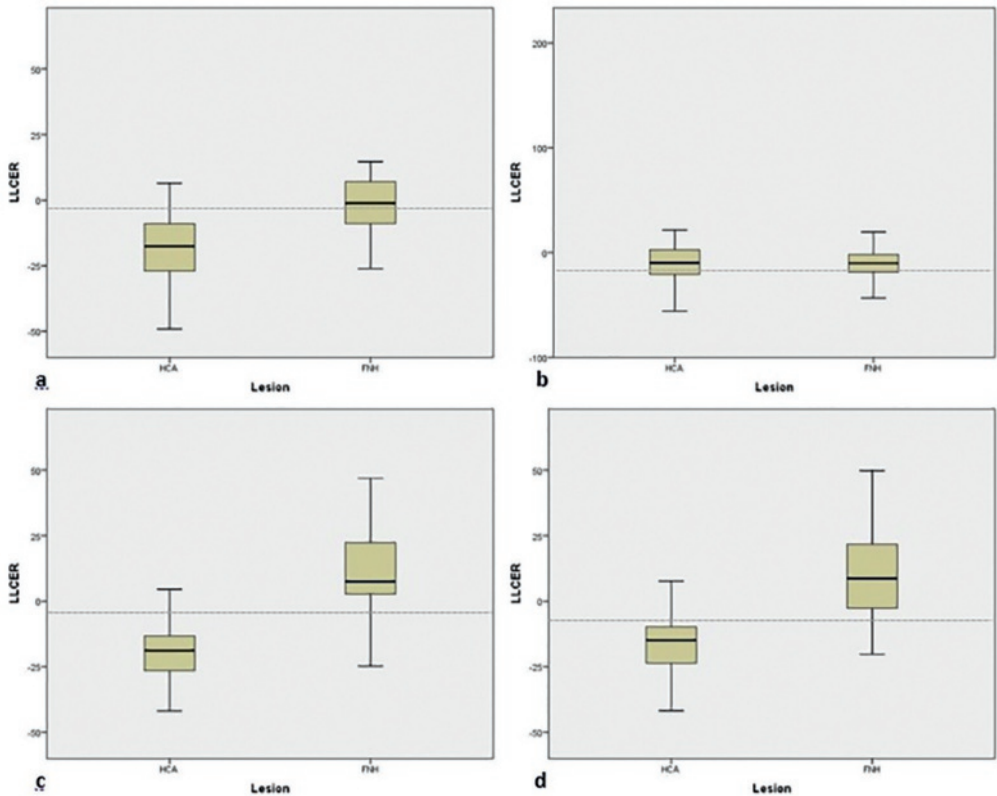
LLCERs were calculated based on readings by two independent readers, both using two different methods, viz. freehand region of interest (ROI) of the entire surface of the lesion (classical method) as performed by Roux et al [1] and a circular ROI (1–3 cm<sup>2</sup>) centered in the region of highest intensity (alternative method).

The interobserver agreement was evaluated using intraclass coefficients (ICC). The Mann-Whitney test was used to determine the difference in LLCER between HCA and FNH. ROC analyses were performed for sensitivities and specificities.

Using the classical method, LLCERs were significantly higher in FNH ( $P < 0.001$ ), but only for observer 1. The area under the curve (AUC) for differentiating HCA from FNH was 0.86 for observer 1 and 0.61 for observer 2. Optimal cutoff values for LLCER and their corresponding sensitivity and specificity were -3.4%, 75.0%, and 89.6% for observer 1, and -15.3%, 75.0%, and 47.9% for observer 2. ICC was fair; 0.58 (0.10–0.81).

Using the alternative method, LLCERs were significantly higher in FNH than HCA ( $P < 0.001$ ) for both observers. The AUC for differentiating HCA and FNH was 0.90 for observer 1 and 0.87 for observer 2. Optimal cutoff values for LLCER and their corresponding sensitivity and specificity were -3.0%, 85.7%, and 91.7% for observer 1, and -6.7%, 82.1%, and 83.3% for observer 2. The ICC was excellent; 0.84 (0.64–0.92).

We could not reproduce optimal cut-offs as previously proposed for the classical method; with our optimal cut-offs, sensitivities and specificities were only mediocre. LLCER assessment using the alternative method was superior, but further studies are necessary to show that the latter method is reproducible.



**Figure 1.** Box plots show the lesion-to-liver contrast enhancement ratio (LLCER) of hepatocellular adenoma (HCA) and focal nodular hyperplasia (FNH) for reader 1 (**a** and **c**) and for reader 2 (**b** and **d**). The two upper plots (**a** and **b**) show the results based on the classical method and the lower two plots show the results based on the alternative method.

The findings of Roux et al. could not be reproduced, which might be due to their lack of a reference standard, but is probably partly due to influences of lesion enhancement in the hepatobiliary phase. Our alternative method only measured regions of maximum enhancement, whereas Roux et al measured the whole surface, which would lead to an underestimation of uptake. We conclude that quantitative analysis of Gd-BOPTA-enhanced hepatobiliary phase MRI is still an ongoing issue of debate.

**Keywords**

hepatocellular adenoma, quantification, FNH, MRI

## References

1. Roux M, Pigneur F, Calderaro J, et al. Differentiation of focal nodular hyperplasia from hepatocellular adenoma: Role of the quantitative analysis of gadobenate dimeglumine-enhanced hepatobiliary phase MRI. *J Magn Reson Imaging* 2015;42(5):1249-1258.
2. Thomeer MG, Willemsen FE, Biermann KK, et al. MRI features of inflammatory hepatocellular adenomas on hepatocyte phase imaging with liver-specific contrast agents. *J Magn Reson Imaging* 2014;39(5):1259-1264.

# Chapter 9

## Intra-patient comparison of the hepatobiliary phase of Gd-BOPTA and Gd-EOB-DTPA in the differentiation of HCA from FNH

Inge JSML Vanhooymissen

**Maarten G Thomeer**

Loes Braun

Bibiche Gest

Sebastiaan van Koeverden

Francois Willemsen

Myriam Hunink

Robert A De Man

Jan N IJzermans

Roy S Dwarkasing

*The first two authors contributed equally to this work*

*Accepted for J Magn Reson Imaging*

## Abstract

**Background:** Current imaging guidelines do not specify the preferred hepatobiliary contrast agent when differentiating hepatocellular adenoma (HCA) from focal nodular hyperplasia (FNH) on MRI.

**Purpose:** To analyze intra-patient differences in the hepatobiliary phase (HBP) after use of both gadobenate dimeglumine (Gd-BOPTA) and gadoxetic acid (Gd-EOB-DTPA)-enhanced MRI to differentiate HCA from FNH.

**Study type:** Retrospective.

**Population:** Patients who underwent both Gd-BOPTA and Gd-EOB-DTPA-enhanced MRI, including 33 patients with 82 lesions (67 HCA; 15 FNH), with a step-down reference standard of pathology, 20% regression, identical appearance to earlier biopsied lesions and stringent imaging findings.

**Field Strength/Sequence:** 1.5T and 3T HBP of Gd-BOPTA and Gd-EOB-DTPA-enhanced MRI, precontrast fat suppressed T1 weighted sequence.

**Assessment:** Signal intensities relative to the surrounding liver in the HBP were assessed by two observers.

**Statistical tests:** Sensitivity and specificity of HCA diagnosis were calculated for both contrast agents. Interobserver agreement was evaluated using Cohen's kappa, differences in degree of certainty for scoring a lesion were calculated by means of the Wilcoxon signed rank test. Differences in signal intensity between Gd-BOPTA and Gd-EOB-DTPA were calculated using McNemars's test.

**Results:** Almost perfect agreement was found between observers for scored signal intensities with both contrast agents. In 30 of the 82 lesions (37%) a difference was observed between contrast agents in the HBP, with Gd-EOB-DTPA proving correct in all but one of the discordant lesions. When distinguishing HCA from FNH, Gd-BOPTA showed a sensitivity of 46% (31/67) and a specificity of 87% (13/15), while the sensitivity and specificity of Gd-EOB-DTPA was 85% (57/67) and 100% (15/15), respectively. A risk of misclassifying HCA as FNH typically occurs for Gd-BOPTA when lesions are intrinsically hyperintense ( $p < 0.005$ ).

**Data conclusion:** The HBP of Gd-EOB-DTPA shows superior accuracy in ruling out HCA in comparison with Gd-BOPTA, especially when the lesion is intrinsically hyperintense on T1-weighted imaging.

### KEYWORDS

Magnetic Resonance Imaging; Liver Cell Adenoma; Focal Nodular Hyperplasia; Guideline; Contrast Media

## Introduction

Focal nodular hyperplasia (FNH) and hepatocellular adenoma (HCA) are both benign lesions of the liver which occur in similar patient populations, yet require very different management [1;2]. FNH is usually managed conservatively without further follow-up, whereas HCA may require surgical treatment or close monitoring due to the known risk of spontaneous hemorrhage [1] and malignant degeneration [2]. A confident diagnosis of FNH versus HCA is therefore important for effective therapeutic decision-making and treatment guidance [3].

MRI is widely used to differentiate the two entities and is based on characteristic imaging features of HCA and FNH. Findings typical for HCA included atoll sign, signs of internal bleeding, cystic areas, and diffuse homogeneous steatosis of the lesion. Findings typical for FNH include an T1-weighted nearly isointense homogenous lesion with T2-weighted hyperintense central scar. However, differentiation can be challenging for lesions without these characteristic features [3-6]. The introduction of liver-specific contrast agents has improved diagnostic accuracy for benign hepatocellular lesions [7-10] and, at present, two contrast agents are in use, [1] gadobenate dimeglumine (Gd-BOPTA, Multihance, Bracco Imaging, Milan, Italy), and [2] gadoxetic acid (Gd-EOB-DTPA, Primovist, Bayer-Schering Pharma, Berlin, Germany).

Initial data has shown that the sensitivity and specificity of the hepatobiliary phase (HBP) for differentiation was over 90% for both contrast agents [8]. However, this data were reported before the introduction of HCA subclassifications. Consequently, earlier studies of accuracy might have been overly optimistic, as was also argued in a recent meta-analysis of Gd-EOB-DTPA accuracy [9]. Because previous telangiectatic FNH were classified as inflammatory HCA, studies before the use of subclassification of HCA are no longer up-to-date. Recently Tse et al. [11] and Grieser et al. [12] used the new classification of HCA as a reference standard and found that inflammatory HCA are mostly hypointense to the surrounding liver in the HBP with the use of Gd-EOB-DTPA. To date, no large studies have been performed using Gd-BOPTA and it therefore remains unclear whether these contrast agents differ regarding the HBP. To avoid confounding due to individual patient characteristics, a potential difference in the HBP can only be analyzed correctly when the same patient is scanned with both agents at different time points.

To date, only one small study has compared both contrast agents in the same patient, but the purpose of that study was the diagnosis of FNH [13]. The goal of the present retrospective study was to review MRI scans of patients who received both liver-specific contrast agents, comparing the value of the HBP in differentiating HCA from FNH and identifying reasons for potential mismatches.

## Materials and Methods

### Study population

The appropriate institutional review board approved this retrospective study and informed consent was waived. We reviewed the MRI data of all patients who underwent MRI for a primary benign liver lesion between June 2008 and December 2016.

The inclusion criteria were: (1) availability of two separate liver MRI examinations with both contrast agents, (2) a diagnostic T1-weighted gradient fat-suppressed HBP scan (slice thickness < 7 mm), (3) a HBP of Gd-EOB-DTPA performed 20 minutes after contrast injection, (4) a HBP of Gd-BOPTA performed at least 1 hour after contrast injection, (5) no known underlying chronic liver disease, (6) a final diagnosis of HCA or FNH (see reference standard), and (7) a lesion size of at least 1 cm.

Exclusion criteria were: (1) other benign liver lesions such as hemangioma, pseudotumor, focal steatosis, etc., and (2) disagreement between readers about typical imaging findings on MRI (see reference standard).

Based on the inclusion criteria, a total of 69 patients were initially reviewed for eligibility to the study, 36 of whom were subsequently excluded. The final study population comprised 33 patients (with 82 lesions) who had undergone both Gd-BOPTA and Gd-EOB-DTPA-enhanced MRI.

All examinations were scanned at our department on a 1.5 or 3 Tesla (T) unit (Philips Medical Systems, Best, The Netherlands, or General Electrics, Signa, Milwaukee, WI; using an identical protocol) or were scanned elsewhere and reviewed at our institute as a second opinion in a tertiary referral center. A comparison of additional information, including dynamic T1-weighted and T2-weighted sequences, was not the focus of this study and was therefore not undertaken. This information was also excluded from the study due to the inconsistency of the scan parameters amongst different patients and scans.

### Imaging evaluation

The HBP of each examination was independently reviewed by two observers: a radiologist specialized in abdominal imaging, with more than twenty years' experience (X1), and a radiologist in his first year of fellowship in abdominal imaging (X2), with five years training in radiology. Both radiologist scored the images separately from each other and started with the Gd-BOPTA scan. After a wash-out period of 1 month they scored the Gd-EOB-DTPA scan. Both were blinded to the final diagnosis, other magnetic resonance (MR) sequences and clinical data. They were not blinded to the contrast agent used when scoring the images. Signal intensities of the lesions in the HBP were assessed and scored as either hypointense or isointense/hyperintense in comparison to the surrounding liver on T1-weighted HBP fat-



suppressed sequences for both contrast agents. A consensus reading was performed in case of disagreement. Both observers graded their degree of certainty (possible, probable, or certain) for labeling a lesion as hypointense or isointense/hyperintense, as hypointense would suggest HCA and isointense/hyperintense would suggest FNH and there might be a difference in certainty when using different contrast agents.

To ensure the independent and blind observation of the hepatobiliary phase, we made sure these two observers (X1 en X2) didn't read other MR sequences. Therefore precontrast fat suppressed T1-weighted images of each examination were independently reviewed by a third and fourth observer: a radiologist specialized in abdominal MR imaging, with more than twenty years' experience (X3), and a radiologist in her 5<sup>th</sup> year of training (X4). All lesions were scored as hypointense, isointense or hyperintense in comparison to the surrounding liver. Both observers were blinded to the final diagnosis, other MR sequences and clinical data. In case of disagreement, a consensus reading was performed.

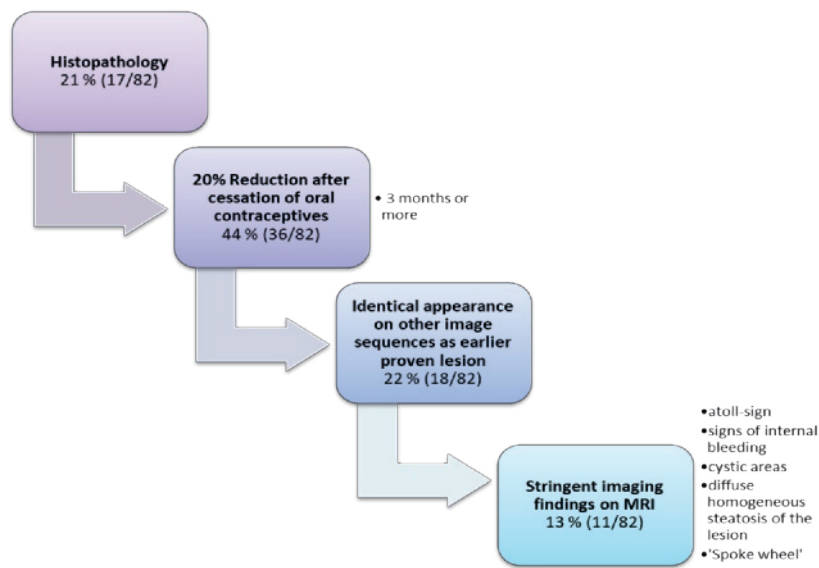
Subclassification of HCA was carried out by a radiologist specialized in abdominal imaging, with more than twenty years' experience (X1). Subclassification was based on a combination of immunohistochemistry and typical imaging findings. Imaging findings diagnostic for inflammatory HCA include the atoll sign, hyperintensity on T2-weighted imaging in combination with arterial hypervascularity, and hyperintensity in the venous phase after Gd-BOPTA injection [4-6]. The diagnostic finding of a steatotic HCA was based on diffuse internal fat in the lesion with homogenous spread [4,6]. These very specific findings were endorsed by a recent international guideline from the EASL [14]. No known specific signs distinguish  $\beta$ -catenin-mutated HCA and therefore all other HCAs were designated as unclassified.

## Reference standard

The reference standard consisted of a step-down process (Figure 1), with the first and most important level of classification based on pathology with immunohistochemistry [15].

The next level included the lesions that were not biopsied or resected. Lesions were classified as HCA when more than 20% regression occurred after at least 3 months cessation of oral contraceptives (while FNHs may vary over time, it has been established that this is not related to cessation of oral contraceptives) [16;17]. To avoid measurement variability, we decided to include only those lesions with prominent decreases in size, defined as a minimum of 20%.

The third level contained the lesions with an identical appearance on other image sequences to earlier adjacent biopsied lesions. Those lesions were classified in accordance with the histopathologic diagnosis of the biopsied or resected lesion. This is common practice, as it is unnecessary and unethical to biopsy all lesions with similar characteristics [18].



**Figure 1.** Step-down process of reference standard.

Finally, the last level contained the remaining lesions which were scored for stringent imaging findings on MRI. These findings were separately scored by two experienced abdominal radiologists (with 25 (X3) and 15 (X5) years of experience, respectively). When disagreement occurred, those lesions were excluded.

Stringent MRI findings typical for HCA included atoll sign, signs of internal bleeding, cystic areas, and diffuse homogeneous steatosis of the lesion [3-6;19]. The atoll sign was first described in 2011, and consists of a T2 hyperintense rim at the periphery of the lesion, with central isointensity [4]. This was typically found for inflammatory HCAs, and was not found in any proven FNHs. To the best of our knowledge, internal bleeding has never been detected in FNH, either with imaging or on pathologic staining [20-22]. The only important differential diagnosis is hepatocellular carcinoma (HCC) [22]. However, other imaging findings were incompatible with HCC, chronic liver disease was absent and lesions did not grow over an interval of at least 4 months on imaging with computed tomography (CT) or MRI. Van Aalten et al. have also shown that internal cystic areas are a reliable tool for differentiation of HCA from FNH [4]. Diffuse internal fat content was also shown to be a very specific sign for steatotic HCAs by several different authors [4-6,23]. Differential diagnoses are mainly HCC and liposarcoma [23]. However, as stated above, none of the lesions showed growth on follow-up. A typical MRI finding for FNH is the presence of a T2-weighted hyperintense central scar (spoke wheel appearance) [18;19].

## Statistical analysis

Sensitivities and specificities were calculated for diagnosing HCA. By definition, FNHs are iso- to hyperintense to the surrounding liver in the HBP, and HCA are routinely hypointense in the HBP [24].

Clustering of the data was adjusted for by using the variance inflation factor [25], and pairing of the data was adjusted for by means of the Durkalski correction [26]. Interobserver agreement was evaluated using Cohen's kappa, using a weighted method based on the squared distance from the diagonal [27]. Differences in degree of certainty for scoring a lesion between the two different contrast agents were calculated for each reader separately by means of the Wilcoxon signed rank test. Differences in signal intensity between Gd-BOPTA and Gd-EOB-DTPA, the influence of the pre-contrast signal intensity on the correctness of the classification, and the discordance in classification between the two contrast agents were all calculated using McNemars's test. A *p*-value of less than 0.05 was considered statistically significant. The analyses were performed in IBM SPSS (version 21) and in R (version 3.3.1) using RStudio (version 0.99.902).

## MRI protocol

The hepatobiliary acquisition consisted typically in the vast majority of the patients of a breath-hold 3D fatsuppressed T1-weighted LAVA sequence (GE Healthcare, Wilconson, WI) both for scans with Gd-BOPTA and Gd-EOB-DTPA. The scan parameters were: TR between 2.7- 3.7 msec, TE 1.1-1.5 msec, field of view (FOV) 40 × 40 cm to 48 × 48 cm (according to body shape), FA = 12-15°, matrix 288 × 256 or higher, NEX = 0.7 or higher, slice thickness 4.0-5.0 mm. We used Gd-EOB-DTPA 0.25 mmol/ml, 1 bottle 10 ml, price 169,10 euro, with dosage of 0.1 ml/kg body weight. And Gd-BOPTA 529 mg/ml, 1 bottle 15 ml, price 39,51 euro, with dosage of 0.1 ml/kg body weight.

## Results

### Study population

The 33 patients included in the study had a total of 82 lesions (67 HCAs and 15 FNHs), and attended our clinic from June 2008 through December 2016. The data variables are shown in Table 1.

The Gd-EOB-DTPA-enhanced MRI was performed first in 79% of the patients (26/33), and the Gd-BOPTA-enhanced MRI was performed first in 21% of the patients (7/33). Most patients received different MRI scans for follow-up after cessation of oral contraceptives. In this way different contrast agents could be used in different sessions. When Gd-EOB-DTPA was the first contrast agent used, from referring centers, mainly in the beginning of the inclusion period, Gd-BOPTA was chosen at our institution for follow-up because of our larger experience with

this agent and the presumed better dynamic phase. When Gd-BOPTA was the first contrast agent used, in several cases Gd-EOB-DTPA was chosen for follow-up mainly to learn from the applicability of latter contrast agent.

In 21% (17/82) of lesions the final diagnosis was based on immunohistochemistry of which 71 % (12/17) HCAs and 29 % (5/17) FNHs (Figure 2), in 44% (36/82) on 20% regression or more after cessation of oral contraceptives for at least 3 months or more (Figure 3), in 22% (18/82) on imaging findings concordant with another histopathology-proven lesion, and in 13% (11/82) on typical imaging findings on MRI.

**Table 1.** Results study population. Time between both MRI examinations = time between included Gd-EOB-DTPA and Gd-BOPTA scan. Follow-up time = total time of follow-up including MRI investigations of the liver without liver specific contrast agent.

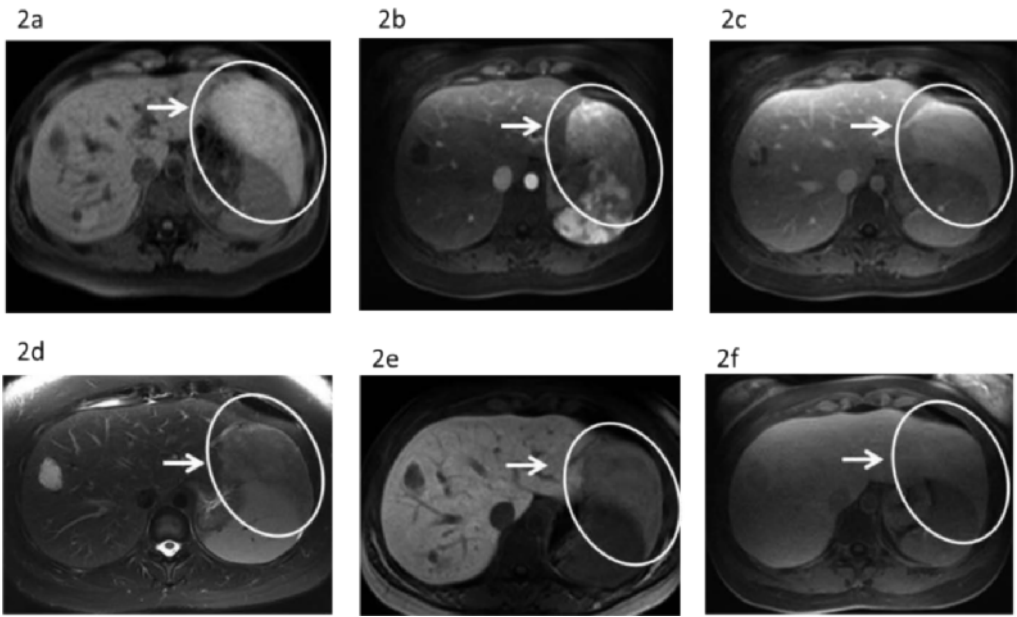
Number of patients	33
Number of lesions	82 (67 HCAs and 15 FNHs)
Gender	All female
Age	Mean: 37 years (age range 22–55 years)
Tumor diameter	40 mm (range 10–100 mm)
Time between both MRI examinations	18 months (min-max: 3–101 months)
Time between arterial phase of Gd-BOPTA and the HBP	1.21 hours (min-max: 1.00–2.20 hours)
Follow-up time	42 months (min-max: 4–101 months)
Gd-EOB-DTPA-enhanced MRI first performed	79 % of the patients (26/33)
Gd-BOPTA-enhanced MRI first performed	21 % of the patients (7/33)

**Interobserver agreement**

Almost perfect agreement was found between the individual readings of both observers for scoring signal intensities in the HBP of all lesions with both contrast agents: Gd-EOB-DTPA: kappa = 0.83 ( $p < 0.001$ ) and Gd-BOPTA: kappa = 0.85 ( $p < 0.001$ ) (Table 2). In only 7% of the lesions (6/82) there was disagreement in the initially recorded signal recording signal intensities, and here consensus reading was carried out. For both observers, no significant differences were found in the degree of certainty for scoring a lesion when using either Gd-EOB-DTPA or Gd-BOPTA. There was a substantial interobserver agreement (kappa = 0.73) for the individual scoring of the precontrast fat-suppressed T1-weighted images, with identical scoring in 77% of the lesions (63/82) (Table 2). For the remaining 23% of the lesions (19/82) consensus was reached.

**Table 2.** Interobserver agreement for pre-contrast sequence and post-contrast sequences with Gd-EOB-DTPA and Gd-BOPTA.

	Kappa	Number of lesions without agreement
Pre-contrast T1W fatsat	0.73 (0.62-0.85)	19 (77%)
HPB of Gd-EOB-DTPA	0.83 (0.70-0.96)	6 (93%)
HPB of Gd-BOPTA	0.85 (0.73-0.96)	6 (93%)

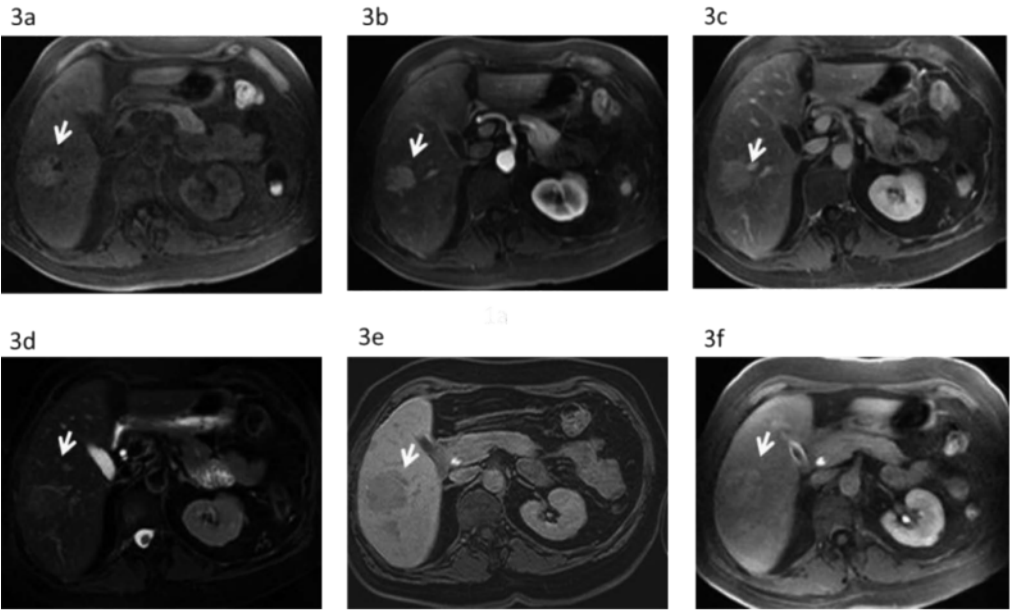


**Figure 2.** A 33-year-old woman with one main lesion (arrow) in the left liver, an inflammatory HCA proven by histopathology. The lesion is slightly hyperintense on T1-weighted imaging without intravenous contrast and with fat saturation (figure 2a). After intravenous contrast injection the lesion is hypervascular in the arterial phase (figure 2b) and slightly hyperintense in the venous phase (figure 2c). Furthermore, the lesion is hyperintense on T2-weighted images (figure 2d), typical of inflammatory HCA [5]. After intravenous injection of Gd-EOB-DTPA the lesion is hypointense in the hepatobiliary phase (figure 2e), whereas the lesion is isointense in the hepatobiliary phase after intravenous injection of Gd-BOPTA (figure 2f).

**Diagnostic Accuracy**

In 30 out of 82 lesions (37%), consisting of 28 HCAs and 2 FNHs, a difference in signal intensity was observed between the two contrast agents in the HBP compared to the surrounding liver. In 29 of 30 discordant lesions, the signal intensity of the lesions in the HBP when using Gd-EOB-DTPA was compatible with the final diagnosis, whereas the HBP of Gd-BOPTA was consistent with the final diagnosis in only one lesion (Figure 4, Figure 5). Of these discordant lesions, the final diagnosis was based on histopathology in 17% (5/30) of cases, on a more

than 20% regression after cessation of oral contraceptives in 63% (19/30), on imaging findings concordant with another histopathology-proven lesion in 13% (4/30), and on typical imaging findings on MRI in 7% (2/30) of cases. Gd-EOB-DTPA was therefore significantly more accurate than Gd-BOPTA ( $p < 0.001$ ).



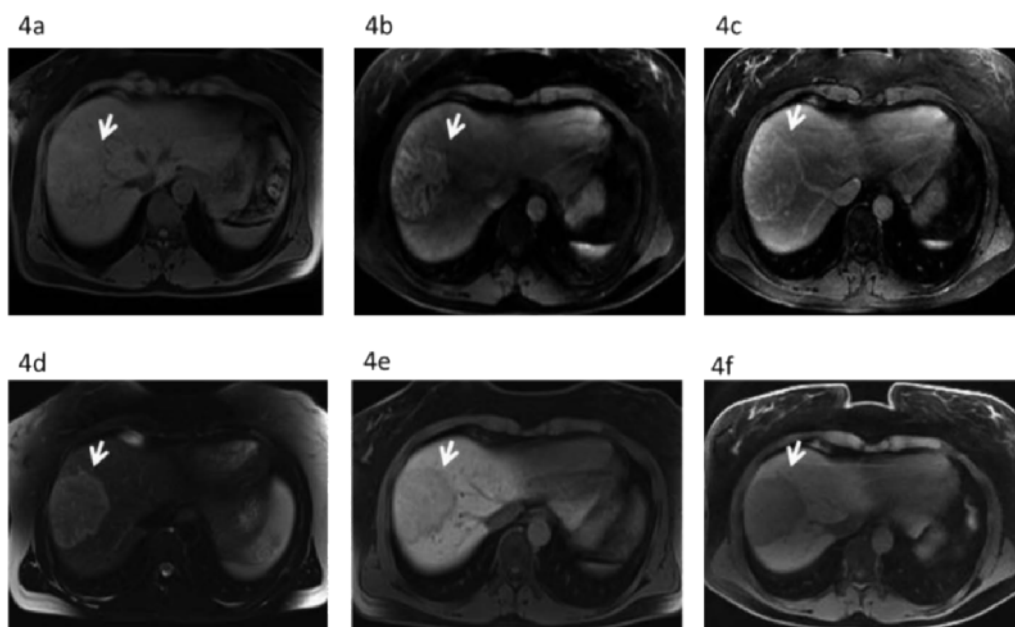
**Figure 3.** A premenopausal 50-year-old woman with an intermediate lesion (arrow) that regressed by more than 20% in maximal diameter over several months after cessation of oral contraceptives, a sign typical of HCA. The lesion is slightly hyperintense on T1-weighted imaging without intravenous contrast and with fat saturation (figure 3a). After intravenous contrast injection the lesion is hypervascular in the arterial phase (figure 3b) and slightly hyperintense in the venous phase (figure 3c). Furthermore, the lesion is slightly hyperintense on T2-weighted images (figure 3d). After intravenous injection of Gd-EOB-DTPA the lesion became hypointense in the hepatobiliary phase (figure 3e), whereas the lesion was hyperintense in the hepatobiliary phase after intravenous injection of Gd-BOPTA (figure 3f).

**Table 3.** Sensitivity and specificity of Gd-BOPTA and Gd-EOB-DTPA in distinguishing HCA from FNH. HCA: hepatocellular adenoma; FNH: focal nodular hyperplasia; AUC: area under the curve.

67 HCA	sensitivity	specificity	AUC
Gd-BOPTA	46 % (31/67)	87 % (13/15)	0.66
Gd-EOB-DTPA	85 % (57/67)	100 % (15/15)	0.93

**Table 4.** Results of the subdivision of HCA and FNH, including the number of congruent and incongruent HBP between Gd-BOPTA and Gd-EOB-DTPA. Note that in all but one incongruent cases, the HBP of Gd-EOB-DTPA provided a correct diagnosis of HCA. pt.nr.: patient number; HCA: hepatocellular adenoma; FNH: focal nodular hyperplasia; HBP: hepatobiliary phase; R: reference step 1 to 4 (see Figure 1).

pt.nr.	HCA	FNH	subclassification HCA	HBP congruent	HBP incongruent	R. step 1	R. step 2	R. step 3	R. step 4
1		1	/	1x					1x
2	4		unclassified	4x			4x		
3	4		inflammatory	4x		1x	3x		
4	2		inflammatory	1x	1x	2x			
5	2		steatotic	2x		1x		1x	
6		1	/		1x	1x			
7	5		unclassified		5x		5x		
8	1		inflammatory	1x		1x			
9	1		inflammatory	1x			1x		
10	1		unclassified	1x			1x		
11	5		inflammatory	1x	4x		5x		
12	1		inflammatory		1x	1x			
13		2	/	2x		1x		1x	
14	2		inflammatory	2x		1x	1x		
15		2	/	2x					2x
16	3		unclassified	3x			3x		
17	3		inflammatory	3x					3x
18		1	/	1x		1x			
19		1	/	1x		1x			
20	2		inflammatory		2x		2x		
21	1		inflammatory		1x				1x
22		4	/	4x		1x		3x	
23	5		inflammatory	5x		1x		4x	
24	1		inflammatory	1x					1x
25	1		inflammatory	1x		1x			
26	5		inflammatory	5x		1x			4x
27	5		2 inflammatory, 3 unclassified	1x	4x		5x		
28	4		unclassified		4x		4x		
29	3		inflammatory	3x			3x		
30	1		inflammatory		1x	1x			
31	5		inflammatory		5x	1x		4x	
32		1	/		1x				1x
33		2	/	2x					2x
67	15			52	30	17	37	13	15



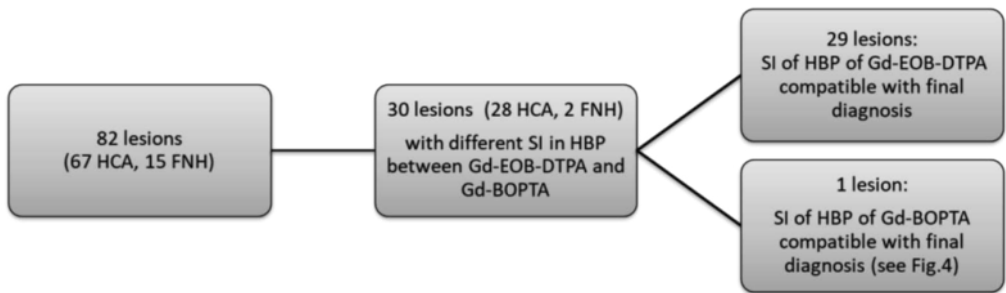
**Figure 4.** A 45-year-old woman with one large lesion (arrow) in the liver that regressed in size by more than 20% after cessation of oral contraceptives, again a characteristic sign of HCA. Moreover, this lesion has the typical aspect of an inflammatory HCA with atoll appearance on T2-weighted imaging (figure 4d). The lesion is isointense on T1-weighted imaging without intravenous contrast and with fat saturation (figure 4a). After intravenous contrast injection the lesion is hypervascular in the arterial phase (figure 4b) and slightly hyperintense, primarily in the periphery in the venous phase (figure 4c). After intravenous injection of Gd-EOB-DTPA the lesion is defined as hypointense by one reader and isointense by the other reader. The lesion was considered isointense in the consensus reading (figure 4e). By contrast, the lesion was uniformly considered hypointense in the hepatobiliary phase after intravenous injection of Gd-BOPTA (figure 4f). This case was the only lesion where a discrepancy between both contrast agents favored Gd-BOPTA.

The diagnostic sensitivity and specificity of Gd-BOPTA in distinguishing HCA from FNH was 46% (31/67) and 87% (13/15), with an area under the curve (AUC) of 0.66, compared to a sensitivity, specificity and AUC for Gd-EOB-DTPA of 85% (57/67), 100% (15/15), and 0.93, respectively (Table 3). Furthermore, HCAs were significantly more often hypointense in the HBP on Gd-EOB-DTPA-enhanced MRI (57/67) versus Gd-BOPTA-enhanced MRI (31/67) ( $p < 0.001$ ) (Figure 6). For FNH lesions, there was no significant difference in signal intensity in the HBP on Gd-EOB-DTPA-enhanced MRI versus Gd-BOPTA-enhanced MRI. When Gd-EOB-DTPA was used, HCA lesions that were hyperintense on pre-contrast T1-weighted images ( $n=39$ ) were classified correctly as HCA in 97% of cases (38/39). When Gd-BOPTA was used, the pre-contrast hyperintense HCA lesions were classified correctly as HCA in 36% of cases (14/39) (Figure 6). Therefore, classification of HCAs that are hyperintense on the pre-contrast sequence is significantly more accurate using Gd-EOB-DTPA compared to Gd-

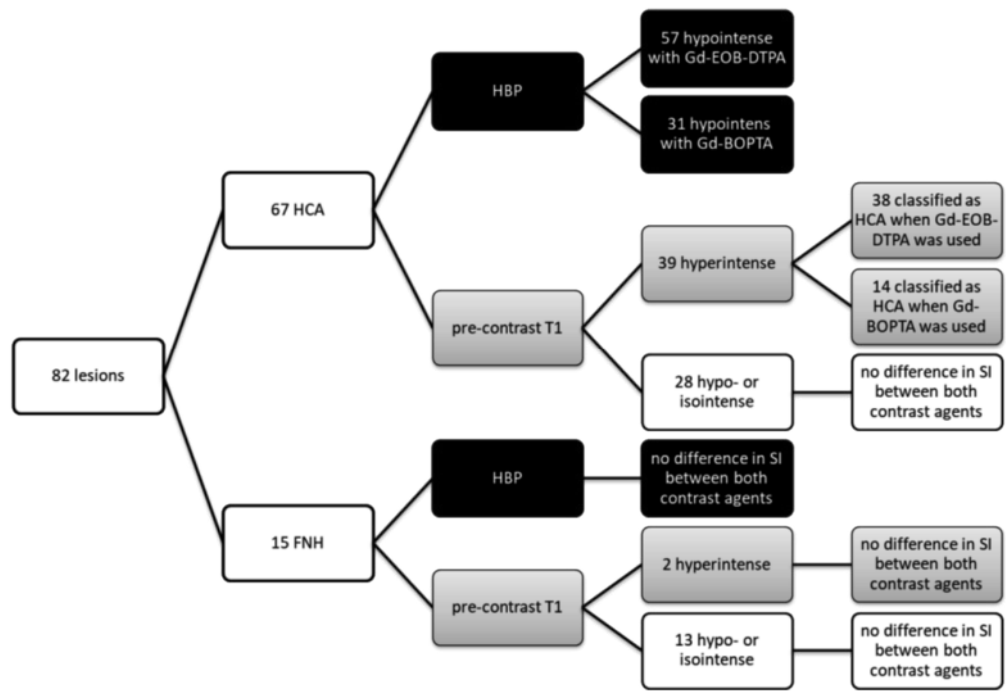


BOPTA ( $p < 0.001$ ), indicating a risk of misclassifying HCA as FNH when Gd-BOPTA is used (Figure 2 and 3). For all FNH lesions, and for HCA lesions that are isointense or hypointense on pre-contrast T1 weighted images, no significant differences in classification based on the pre-contrast T1-weighted images were found for either Gd-BOPTA or Gd-EOB-DTPA.

Subclassifications of HCA are presented in Table 4.



**Figure 5.** Flowchart of the 30 lesions where a difference in signal intensity was observed between the two contrast agents in the hepatobiliary phase (HBP) compared to the surrounding liver. SI: signal intensity.



**Figure 6.** Flowchart of the results of both contrast agents after scoring the hepatobiliary phase (HBP). And the results of the lesions which are hyperintense on the pre-contrast T1 weighted images. SI: signal intensity.

## Discussion

Using intra-patient comparisons, this study shows that Gd-EOB-DTPA and Gd-BOPTA are not equivalent in the HBP when the goal is distinguishing HCA from FNH. We found that Gd-BOPTA was too inconsistent to reliably exclude HCA. This means that if the lesion is iso/hyperintense to the surrounding liver in the HBP, Gd-BOPTA cannot differentiate between the 2 entities with significant reliability. When HCAs are hyperintense on pre-contrast T1-weighted images, incorrect diagnosis occurs significantly more often when using Gd-BOPTA. However, if the lesion is hypointense to the surrounding liver in the HBP, both contrast agents can reliably diagnose HCA.

Although Gd-BOPTA and Gd-EOB-DTPA are both liver-specific contrast agents, there is a well-established difference between the two in terms of hepatocyte uptake and excretion by the biliary system. While Gd-EOB-DTPA has an excretion level of 50%, the excretion level of Gd-BOPTA (5%) results in a relatively weak liver signal intensity and biliary tree enhancement [24]. This interpretation is supported by the fact that significantly more HCAs were isointense/hyperintense in the HBP, but at the same time hyperintense on precontrast T1-weighted imaging. The combination of intrinsic hyperintensity of HCAs and weak uptake of Gd-BOPTA in the liver probably gives a false impression of uptake in the lesion, resulting in misclassification of HCA as FNH.

Published reports on the use of the Gd-BOPTA HBP in the diagnosis of HCA are relatively sparse [8;28;29], and no large data sets are available that used stringent reference standards. Moreover, when a pathology reference standard was used, it rarely included immunohistochemistry. Published reports on the use of the Gd-EOB-DTPA HBP are much more common, and two separate meta-analyses reported high accuracy, both in terms of sensitivity and specificity [30;31]. Although these studies also discussed some mixed results, adequately-performed studies based on histochemistry were almost uniformly positive.

The possibility that variation in MRI quality or reading experience could explain the observed differences between Gd-BOPTA and Gd-EOB-DTPA in our study is lessened by the high degree of certainty of both readers with both contrast agents and the rare occurrence of disagreement.

Since the establishment of the new HCA classification by the Bordeaux group, including the subtypes hepatocyte nuclear factor 1 $\alpha$  (HNF1 $\alpha$ )-mutated,  $\beta$ -catenin-mutated, inflammatory HCA, and unclassified HCA [32;33], mounting evidence in the literature indicates that the inflammatory and  $\beta$ -catenin HCAs are the primary subtypes that need to be ruled out. Although not all included lesions were confirmed by pathology, most of discrepancies occurred in inflammatory subtype HCAs. Furthermore, in addition to the fact that these lesions can be intrinsically hyperintense, they may take up contrast in the HBP to a certain degree [11,19,34].

This characteristic might go some way to explaining why inflammatory HCAs in our study population were more often iso- to hyperintense in the HBP using Gd-BOPTA.

The major limitations of the study were the retrospective design and the lack of pathologic confirmation of all lesions. However, with the additional support of typical and stringent clinical and MRI findings classified by experienced radiologists, our results appear valid and consistent with available literature [30].

Another limitation was likely selection bias. Firstly, FNH was underrepresented in our results, probably because diagnosis of FNH is primarily based on imaging, especially in experienced centers. As recently stated in a clinical guideline by Vilgrain et al., FNH is mainly diagnosed by radiology, with biopsy only warranted in difficult cases [35]. Secondly, several patients had multiple lesions, which might have caused selection bias. However, we performed a statistical correction of the data by means of a variance inflation factor and the Durkalski correction to minimize this effect. Thirdly, when a lesion showed a typical sign of HCA but was hyper/isointense in the HBP with Gd-BOPTA, an additional Gd-EOB-DTPA scan was performed to confirm the diagnosis of HCA. Nine of the 33 patients (27%) first received a scan with Gd-BOPTA, but the subsequent MRI with Gd-EOB-DTPA was performed in only two of these patients, due to inconsistencies between the HBP (Gd-BOPTA) and other sequences concerning final diagnosis. In these situations, difficult and incongruent Gd-BOPTA scans might be overrepresented in our study population. Of note, in none of the cases was a scan with Gd-BOPTA envisaged following inconsistencies between the HBP of Gd-EOB-DTPA and other sequences concerning final diagnosis.

One might argue that, besides the HBP, imaging diagnosis of these benign entities is generally based on a combination of all imaging parameters. In most circumstances, final diagnoses can be reached based only on typical findings [4-6;36]. Although the aim of our study was to specifically clarify the utility of the HBP in establishing a final diagnosis, future studies could address the question of the relative importance of the HBP compared to other imaging parameters in typical cases. Nonetheless, lesions that do not show typical imaging findings are not uncommon and therefore require confirmation on the HBP. In our opinion this issue may be even more prominent in less experienced centers, where the use of the HBP of Gd-BOPTA could result in misdiagnosis, with potentially important clinical consequences.

Although the biases were substantial, we found that our findings (in all but one incongruent lesion Gd-EOB-DTPA was finally correct) are still significant enough to change our policy to the preferential use Gd-EOB-DTPA. Even though a prospective study using both contrast agents in the same patients would help confirm our current findings, we would have considerable ethical difficulty with this approach in the light of current discussion concerning linear contrast agents and deposition in the brain [36].

Which contrast agent is currently preferred when liver MRI is proposed? No advice is given concerning a preferred MRI contrast agent in the recently published (2016) EASL clinical practice guidelines on the management of benign liver tumors [35]. According to the ESGAR consensus statement on liver-specific contrast agents (2016), there are no data indicating diagnostic superiority of one agent over another [37], but it was also noted that few publications have compared Gd-EOB-DTPA and Gd-BOPTA and that no differentiation was made between HCA and FNH in those that did [13;38;39].

In light of the new insights provided by our study (keeping in mind the possible biases), we suggest that Gd-EOB-DTPA is preferable when the question is solely differentiation of HCA and FNH. This type of situation may arise due to, for instance, a combination of age, sex and previous imaging findings (i.e. multiphase CT-scan) in which the pre-MRI probability for HCA or FNH is very high.

When the clinical request is to diagnose an atypical liver lesion, a non-specific contrast agent might be chosen first, mainly because these agents have a potentially better arterial and venous dynamic phase [19]. Wash-out in the venous phase is an important feature that adds to a suspicion of malignancy, such as in case of hypervascular metastases and hepatocellular carcinoma in a non-cirrhotic liver (40). In a second phase, with lesions showing only benign features on the first MRI, an additional MRI with Gd-EOB-DTPA could be proposed as a problem solver.

In conclusion, Gd-EOB-DTPA and Gd-BOPTA show significant differences in the HBP. HBP findings based on Gd-BOPTA should be avoided, especially if the lesion is iso- to hyperintense to the surrounding liver in this scan phase. In general, Gd-EOB-DTPA appears more accurate when the goal is differentiation of HCA from FNH. However, further confirmation in larger series is needed.

## References

1. van Aalten SM, de Man RA, JN IJ, Terkivatan T. Systematic review of haemorrhage and rupture of hepatocellular adenomas. *Br J Surg* 2012;99(7):911-916.
2. Nagorney DM. Benign hepatic tumors: focal nodular hyperplasia and hepatocellular adenoma. *World J Surg* 1995;19(1):13-18.
3. Hussain SM, van den Bos IC, Dwarkasing RS, Kuiper JW, den Hollander J. Hepatocellular adenoma: findings at state-of-the-art magnetic resonance imaging, ultrasound, computed tomography and pathologic analysis. *Eur Radiol* 2006;16(9):1873-1886.
4. van Aalten SM, Thomeer MG, Terkivatan T, et al. Hepatocellular adenomas: correlation of MR imaging findings with pathologic subtype classification. *Radiology* 2011;261(1):172-181.
5. Ronot M, Bahrami S, Calderaro J, et al. Hepatocellular adenomas: accuracy of magnetic resonance imaging and liver biopsy in subtype classification. *Hepatology* 2011;53(4):1182-1191.
6. Laumonier H, Bioulac-Sage P, Laurent C, Zucman-Rossi J, Balabaud C, Trillaud H. Hepatocellular adenomas: magnetic resonance imaging features as a function of molecular pathological classification. *Hepatology* 2008;48(3):808-818.
7. Grazioli L, Bondioni MP, Haradome H, et al. Hepatocellular adenoma and focal nodular hyperplasia: value of gadoxetic acid-enhanced MR imaging in differential diagnosis. *Radiology* 2012;262(2):520-529.
8. Grazioli L, Morana G, Kirchin MA, Schneider G. Accurate differentiation of focal nodular hyperplasia from hepatic adenoma at gadobenate dimeglumine-enhanced MR imaging: prospective study. *Radiology* 2005;236(1):166-177.
9. Suh CH, Kim KW, Kim GY, Shin YM, Kim PN, Park SH. The diagnostic value of Gd-EOB-DTPA-MRI for the diagnosis of focal nodular hyperplasia: a systematic review and meta-analysis. *Eur Radiol* 2015;25(4):950-960.
10. Bieze M, van den Esschert JW, Nio CY, et al. Diagnostic accuracy of MRI in differentiating hepatocellular adenoma from focal nodular hyperplasia: prospective study of the additional value of gadoxetate disodium. *AJR Am J Roentgenol* 2012;199(1):26-34.
11. Tse JR, Naini BV, Lu DS, Raman SS. Qualitative and Quantitative Gadoxetic Acid-enhanced MR Imaging Helps Subtype Hepatocellular Adenomas. *Radiology* 2016;279(1):118-127.
12. Grieser C, Steffen IG, Kramme IB, et al. Gadoxetic acid enhanced MRI for differentiation of FNH and HCA: a single centre experience. *Eur Radiol* 2014;24(6):1339-1348.
13. Gupta RT, Iseman CM, Leyendecker JR, Shyknevsky I, Merkle EM, Taouli B. Diagnosis of focal nodular hyperplasia with MRI: multicenter retrospective study comparing gadobenate dimeglumine to gadoxetate disodium. *AJR Am J Roentgenol* 2012;199(1):35-43.
14. European Association for the Study of the L. EASL Clinical Practice Guidelines on the management of benign liver tumours. *Journal of hepatology* 2016;65(2):386-398.
15. van Aalten SM, Verheij J, Terkivatan T, Dwarkasing RS, de Man RA, IJzermans JN. Validation of a liver adenoma classification system in a tertiary referral centre: implications for clinical practice. *Journal of hepatology* 2011;55(1):120-125.
16. Rabe T, Feldmann K, Grunwald K, Runnebaum B. Liver tumours in women on oral contraceptives. *Lancet* 1994;344(8936):1568-1569.
17. Kapp N, Curtis KM. Hormonal contraceptive use among women with liver tumors: a systematic review. *Contraception* 2009;80(4):387-390.
18. Terkivatan T, van den Bos IC, Hussain SM, Wielopolski PA, de Man RA, JN IJ. Focal nodular hyperplasia: lesion characteristics on state-of-the-art MRI including dynamic gadolinium-enhanced and superparamagnetic iron-oxide-uptake sequences in a prospective study. *J Magn Reson Imaging* 2006;24(4):864-872.
19. Thomeer MG, ME EB, de Lussanet Q, et al. Genotype-phenotype correlations in hepatocellular adenoma: an update of MRI findings. *Diagn Interv Radiol* 2014;20(3):193-199.

20. Balabaud C, Al-Rabih WR, Chen PJ, et al. Focal Nodular Hyperplasia and Hepatocellular Adenoma around the World Viewed through the Scope of the Immunopathological Classification. *Int J Hepatol* 2013;2013:268625.
21. Hussain SM, Terkivatan T, Zondervan PE, et al. Focal nodular hyperplasia: findings at state-of-the-art MR imaging, US, CT, and pathologic analysis. *Radiographics* 2004;24(1):3-17; discussion 18-19.
22. Li RK, Zeng MS, Rao SX, et al. Using a 2D multibreath-hold susceptibility-weighted imaging to visualize intratumoral hemorrhage of hepatocellular carcinoma at 3T MRI: correlation with pathology. *J Magn Reson Imaging* 2012;36(4):900-906.
23. Costa AF, Thipphavong S, Arnason T, Stueck AE, Clarke SE. Fat-Containing Liver Lesions on Imaging: Detection and Differential Diagnosis. *AJR Am J Roentgenol* 2017:1-10.
24. Seale MK, Catalano OA, Saini S, Hahn PF, Sahani DV. Hepatobiliary-specific MR contrast agents: role in imaging the liver and biliary tree. *Radiographics* 2009;29(6):1725-1748.
25. Genders TS, Spronk S, Stijnen T, Steyerberg EW, Lesaffre E, Hunink MG. Methods for calculating sensitivity and specificity of clustered data: a tutorial. *Radiology* 2012;265(3):910-916.
26. Durkalski VL, Palesch YY, Lipsitz SR, Rust PF. Analysis of clustered matched-pair data. *Stat Med* 2003;22(15):2417-2428.
27. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33(1):159-174.
28. Thomeer MG, Willemssen FE, Biermann KK, et al. MRI features of inflammatory hepatocellular adenomas on hepatocyte phase imaging with liver-specific contrast agents. *J Magn Reson Imaging* 2014;39(5):1259-1264.
29. Morana G, Grazioli L, Kirchin MA, et al. Solid hypervascular liver lesions: accurate identification of true benign lesions on enhanced dynamic and hepatobiliary phase magnetic resonance imaging after gadobenate dimeglumine administration. *Invest Radiol* 2011;46(4):225-239.
30. McInnes MD, Hibbert RM, Inacio JR, Schieda N. Focal Nodular Hyperplasia and Hepatocellular Adenoma: Accuracy of Gadoxetic Acid-enhanced MR Imaging--A Systematic Review. *Radiology* 2015;277(2):413-423.
31. Guo Y, Li W, Xie Z, et al. Diagnostic Value of Gd-EOB-DTPA-MRI for Hepatocellular Adenoma: A Meta-Analysis. *J Cancer* 2017;8(7):1301-1310.
32. Bioulac-Sage P, Blanc JF, Rebouissou S, Balabaud C, Zucman-Rossi J. Genotype phenotype classification of hepatocellular adenoma. *World journal of gastroenterology: WJG* 2007;13(19):2649-2654.
33. Zucman-Rossi J, Jeannot E, Nhieu JT, et al. Genotype-phenotype correlation in hepatocellular adenoma: new classification and relationship with HCC. *Hepatology* 2006;43(3):515-524.
34. Ba-Ssalamah A, Antunes C, Feier D, et al. Morphologic and Molecular Features of Hepatocellular Adenoma with Gadoxetic Acid-enhanced MR Imaging. *Radiology* 2015;277(1):104-113.
35. European Association for the Study of the Liver . Electronic address eee. EASL Clinical Practice Guidelines on the management of benign liver tumours. *Journal of hepatology* 2016;65(2):386-398.
36. Runge VM. Critical Questions Regarding Gadolinium Deposition in the Brain and Body After Injections of the Gadolinium-Based Contrast Agents, Safety, and Clinical Recommendations in Consideration of the EMA's Pharmacovigilance and Risk Assessment Committee Recommendation for Suspension of the Marketing Authorizations for 4 Linear Agents. *Invest Radiol* 2017;52(6):317-323.
37. Neri E, Bali MA, Ba-Ssalamah A, et al. ESGAR consensus statement on liver MR imaging and clinical use of liver-specific contrast agents. *Eur Radiol* 2016;26(4):921-931.
38. Karam AR, Shankar S, Surapaneni P, Kim YH, Hussain S. Focal nodular hyperplasia: central scar enhancement pattern using Gadoxetate Disodium. *J Magn Reson Imaging* 2010;32(2):341-344.
39. Park Y, Kim SH, Kim SH, et al. Gadoxetic acid (Gd-EOB-DTPA)-enhanced MRI versus gadobenate dimeglumine (Gd-BOPTA)-enhanced MRI for preoperatively detecting hepatocellular carcinoma: an initial experience. *Korean J Radiol* 2010;11(4):433-440.
40. Colombo M. Diagnosis of liver nodules within and outside screening programs. *Ann Hepatol* 2015;14(3):304-309.

# Chapter 10

## Hepatocellular adenoma: when and how to treat? Update of current evidence

Maarten G Thomeer

Mirelle Broker

Joanne Verheij

Michael Doukas

Turkan Terkivatan

Diederick Bijdevaate

Robert A De Man

Adriaan Moelker

Jan N IJzermans

*The first two authors contributed equally to this work*

## Abstract

Hepatocellular adenoma (HCA) is a rare, benign liver tumor. Discovery of this tumor is usually as an incidental finding, correlated with the use of oral contraceptives, or pregnancy. Treatment options have focused on conservative management for the straightforward, smaller lesions (<5 cm), with resection preferred for larger lesions (> 5 cm) that pose a greater risk of hemorrhage or malignant progression. In recent years, a new molecular subclassification of HCA has been proposed, associated with characteristic morphological features and loss or increased expression of immunohistochemical markers. This subclassification could possibly provide considerable benefits in terms of patient stratification, and the selection of treatment options. In this review we discuss the decision-making processes and associated risk analyses that should be made based on lesion size, and subtype. The usefulness of this subclassification system in terms of the procedures instigated as part of the diagnostic workup of a suspected HCA will be outlined, and suitable treatment schemes proposed.

## Keywords

adenoma, immunohistochemistry, liver, MRI



## Introduction

Hepatocellular adenomas (HCAs) are an uncommon, solid, benign tumor of the liver, with an estimated incidence of 3–4 per 100,000 women [Bioulac-Sage et al. 2010]; this frequency is based on population research including women using oral contraceptives (OCs) [Baum et al. 1973]. A causal role for hormone activity in HCA growth is suggested by data linking adenoma regression to the cessation of OC use, and growth associated with pregnancy [Cobey and Salem, 2004].

Typically, HCAs are treated conservatively, with patients advised to avoid oral contraception. The risks of growth and rupture of HCAs during pregnancy has to be underlined, especially in larger HCAs. Tumor progression, suggested by internal bleeding and malignant transformation, necessitates a more aggressive therapeutic approach, with lesions larger than 5 cm considered as the primary risk factor [Marrero et al. 2014]. The introduction of a new subclassification system for HCA has been suggested to help clinicians to stratify patients according to imaging criteria, expression of associated immunohistochemical markers or molecular findings. These data may influence the treatment selected [Marrero et al. 2014] since certain subtypes of HCA pose a greater risk of progression to hepatocellular carcinoma (HCC) than others. For example, a subtype of HCA defined by the reduced expression of liver fatty acid-binding protein (LFABP) ordinarily indicates a subtype with a less aggressive course and a tendency towards a benign phenotype.

Based on the recent literature, we describe the impact of this newly instigated HCA subclassification, and discuss whether this knowledge, combined with imaging data, improves our risk analyses for patients with HCA. Furthermore, we outline the different therapeutic options indicated by each HCA subtype.

## The Bordeaux classification of HCA

In recent years, four distinct subtypes of HCA have been recognized: inflammatory HCA (40–50%, IHCA), HNF1A-mutated HCA (30–40%, H-HCA),  $\beta$ -catenin activated HCA (10–15% b-HCA), and unclassified HCA (10–25%, UHCA) [Nault et al. 2013]. In these different subtypes, several genetic mutations are identified, causing (benign) proliferation of hepatocytes and in some HCA, malignant transformation [Pilati et al. 2014].

Patients presenting with IHCA demonstrate both serum, and lesional indicators of an active inflammatory response. In these lesions, increased expression is seen of the markers serum amyloid A and C-reactive protein, both classic indicators of the acute phase response [Bioulac-Sage et al. 2007a]. Patients within this HCA category frequently demonstrate increased body weight, and a high alcohol intake [Bioulac-Sage et al. 2007b, 2009; Paradis et al. 2007]. In approximately 10–20% of these lesions, a  $\beta$ -catenin mutation is found [van Aalten et al. 2011b].

A second subtype of HCA, H-HCA, is characterized by a downregulation of the LFABP; this phenotype, which is not apparent in the other HCA subtypes, rarely leads to malignant progression [Zucman-Rossi et al. 2006].

A second subtype of HCA, H-HCA, is characterized by a downregulation of the LFABP; this phenotype, which is not apparent in the other HCA subtypes, rarely leads to malignant progression [Zucman-Rossi et al. 2006].

Subtype b-HCA is typified by activating mutations of  $\beta$ -catenin that resist phosphorylation-mediated down-regulation by the GSKB/APC/ AXIN complex [Nault et al. 2013]. Particularly the exon 3 mutation of  $\beta$ -catenin plays a significant role in malignant progression in contrast to exon 7/8 mutations [Pilati et al. 2014]. The result is an accumulation of nuclear  $\beta$ -catenin which, combined with deletion of APC, favors progression to HCC [Nault et al. 2013]. The comparatively small number of  $\beta$ -catenin positive nuclei can lead to this phenotype being overlooked in small biopsies [van Aalten et al. 2011b]. The b-HCA subtype also demonstrates an overexpression of glutamate-ammonia ligase (GLUL) encoding glutamine synthase (GS), which can be used as a sensitive diagnostic biomarker for this subtype [van Aalten et al. 2011b].

The final subtype, UHCA, is not yet defined by any specific genetic mutation, but is instead characterized by various histologic criteria that are unusual in the other subtypes; the underlying pathogenesis of this subtype remains unclear [Blanc et al. 2015].

## **Magnetic resonance imaging of the different subtypes of HCA**

The primary differential diagnosis for HCA is focal nodular hyperplasia (FNH). If in doubt, a biopsy should be taken, especially for larger lesions, as the clinical management will differ for either pathology. In most cases these diagnoses can be differentiated according to signal intensity and dynamic vascular patterns after intravenous aspecific gadolinium injection [(conventional magnetic resonance imaging (MRI)) [van Aalten et al. 2011]. Different patterns can be used for confident diagnosis as proposed by Thomeer and colleagues [Thomeer et al. 2014a].

In more challenging cases, specific hepatobiliary contrast agents can be used. Currently there are two agents available, gadobenate dimeglumine, and gadoxetate disodium [Grazioli et al. 2013; Thomeer et al. 2014a; McInnes et al. 2015]. If the lesion turns hypointense to the surrounding liver in the hepatobiliary phase, FNH can be excluded in most cases. If the lesion becomes iso- to hyperintense the differential diagnosis is FNH or in exceptional cases HCC. However, it should be noted that IHCA can also be isointense in the hepatobiliary phase [Agarwal et al. 2014; Thomeer et al. 2014b]. This might be explained by the presence of internal bile duct proliferation, previously thought to be only visible in FNH. This diagnostic pitfall can be visualized when using gadobenate dimeglumine [Thomeer et al. 2014], or gadoxetate

disodium [Agarwal et al. 2014a]. A recent systematic review about the value of gadoxetate disodium has shown that apart from this pitfall, adequate differentiation is possible in most cases [McInnes et al. 2015]. It was reported that the hepatobiliary phase has a sensitivity of 91–100% and a specificity of 87–100% for differentiating HCA from FNH. In the largest study this was only seen in 13% of cases [Bieze et al. 2012].

In conclusion, in the vast majority lesions can easily be differentiated based on a combination of typical findings on conventional MRI and features on hepatobiliary phase MRI.

Some typical MRI features allow us to discriminate different subtypes of HCA (Table 1): IHCA can be hyperintense on T2-weighted images, with persistent enhancement on delayed imaging in the venous phase [Laumonier et al. 2008]. Ronot and colleagues validated this feature as being highly specific for IHCA, with a sensitivity of 82% [28/34, confidence interval (CI) 65–93%], and an optimal specificity of 100% (12/12, CI 75–100%) [Ronot et al. 2011]. Another diagnostic indicator for IHCA is the atoll-sign [van Aalten et al. 2011a], a hyperintense rim on T2-weighted images at the periphery of the lesion (resembling an atoll) that is enhanced late in the venous phase. This feature is present in 27% of IHCAs in this study [van Aalten et al. 2011a].

Whilst H-HCAs are typically characterized by a large amount of aberrant fat which can be readily appreciated by out-of-phase imaging, or on a fatsaturated T1-weighted image [Laumonier et al. 2008; van Aalten et al. 2011a], van Aalten and colleagues failed to detect fat by MRI for as many as 22% of cases (2/9) [van Aalten et al. 2011a]. Ronot and colleagues validated the diagnostic feature of diffuse and homogeneous signal dropout on outof- phase T1 weighted imaging, with a reported sensitivity of 90% (10/11, CI 58–99), and specificity of 88% (32/36, CI 73–96) [Ronot et al. 2011]. The main drawback of this marker is that diffuse intralesional steatosis may also be present in up to 11% (4/34) of IHCAs [Ronot et al. 2011], although, according to the authors, this does not represent a major pitfall as fat is usually distributed heterogeneously within IHCAs (Figure 1).

The MRI features of b-HCA are not well defined, principally because these lesions are comparatively rare. Van Aalten and colleagues reported poorly delimited, high-signal intensity areas, to be typical of this subtype (5/7, 71%), but additional investigations are warranted [van Aalten et al. 2011a]. Table 2 shows the various MRI features reported in the literature for b-HCA, although, where reported, these features are inconsistent. Despite significant numbers of false negatives, the specificity of these MRI features is very high, leading us to conclude that if any one of these signs are present, a diagnosis of the corresponding MRI subtype can be made with some certainty. Larger datasets will be needed to determine the true value of MRI in HCA imaging for all subtypes; currently, this technique is of most use in evaluating prognosis.

Reviewing the known complications

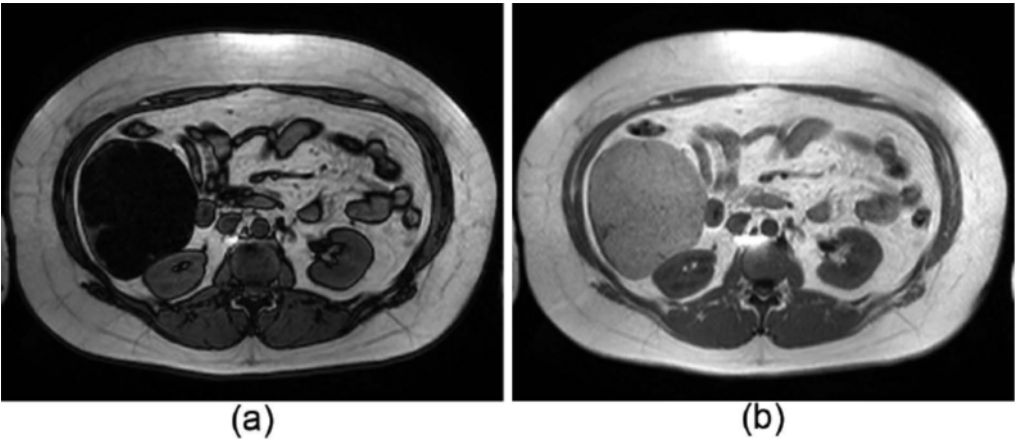
*Intralesional bleeding*

On reviewing the recent literature, van Aalten and colleagues detected evidence of hemorrhage in 27.2% of all patients (315/1160) with one or more HCAs, giving a 15.8% chance of hemorrhage for every HCA (118/748) [van Aalten et al. 2012]. Acute rupture and intraperitoneal bleeding were reported in 17.5% of patients. A size for the smallest HCAs showing hemorrhage was reported for 13 of the 28 articles reviewed; hemorrhage generally arose in the larger lesions (>5 cm), although smaller lesions could also bleed (Table 3), albeit at much lower rates. These data should be interpreted with caution, as only the resected cases were included. The actual chance of bleeding in the different subtypes is likely to be significantly lower. The risk of bleeding was inconsistent across the subtypes of HCA: IHCA showed a higher propensity for macroscopic hemorrhage (30%), than H-HCA (8%) [Dokmak et al. 2009] which can presumably be attributed to the larger number of venous structures, or telangiectatic changes in this subtype.

**Table 1.** Typical MRI findings according to subtypes of HCA.

Subtype	Most typical MRI signs
IHCA	Hyperintense on T2-weighted images, with persistent enhancement in the venous phase; atoll sign
H-HCA	Diffuse and homogenous fat deposition (Figure 1)
b-HCA	(Vaguely demarcated scar)
UHCA	No typical sign

b-HCA,  $\beta$ -catenin activated HCA; HCA, hepatocellular adenoma; H-HCA, HNF1A-mutated HCA; IHCA, inflammatory HCA; MRI, magnetic resonance imaging; UHCA, unclassified HCA.



**Figure 1.** In- and out-of-phase MRI of a typical case of histochemistry proved H-HCA which was resected. Note the diffuse and homogenous signal drop-off on the out-of-phase image (a) versus the in-phase image (b). This correlates with diffuse intralesional fat identified by histology. MRI differentiation between H-HCA and IHCA would not be possible when the signal drop is more heterogeneous. HCA, hepatocellular adenoma; H-HCA, HNF1A-mutated HCA; IHCA, inflammatory HCA; MRI, magnetic resonance imaging.

Although there may be a difference in prevalence of internal bleeding, all subtypes bear this intrinsic risk [Laumonier et al. 2008; Dokmak et al. 2009; Ronot et al. 2011; van Aalten et al. 2011], which diminishes the utility of subtype classification in terms of the clinical management of this risk. Furthermore, more data are needed to prove any correlate between reduced bleeding and the H-HCA subtype.

Bieze and colleagues described a series of 45 patients with 195 lesions. In this cohort, there was a tendency to an enhanced risk of bleeding when the lesion was located in the left lateral liver (11/32 versus 31/163 in other regions), and showed exophytic growth (16/24 versus 9/82) [Bieze et al. 2014] (Figure 1). The latter phenomenon is probably due to the subcapsular location, with no intrinsic capsule, and minimal surrounding liver with which to prevent rupture of the hematoma into the abdominal cavity. However, no other data are available to support this theory, and preventive treatment in these cases does not appear to be warranted.

As for the clinical application of these findings, there is no evidence that supports the use of subtype classification in the stratification and management of individual patients. Moreover, size still remains the most important marker to predict those at risk for larger bleeding in follow up.

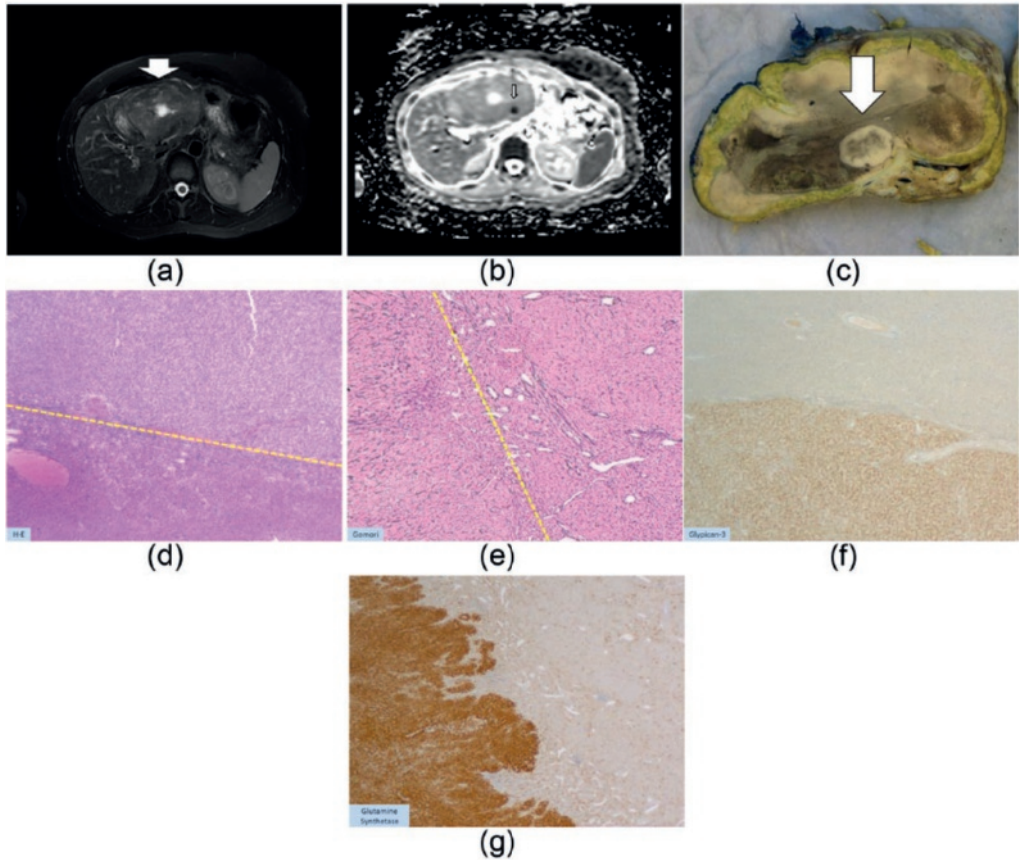
**Table 2.** Recently published b-HCAs with their typical characteristics defined by MRI. Note the low prevalence in the literature of MRI data, with inconsistent findings.

Authors	Year of publication	Number of $\beta$ -catenin HCA	MRI findings
[Van Aalten et al. 2011]		7	Vaguely defined scar on T2W sequences (3 cases)
[Laumonier et al. 2008]	2008	5	Marked hyperintensity on T2W sequences and persistent delayed enhancement (3 cases) Isointensity on T2W sequences, with strong arterial enhancement and delayed washout (2 cases)
[Yoneda et al. 2012]	2012	1	Vaguely defined scar on T2W sequences

b-HCA,  $\beta$ -catenin positive HCA; HCA, hepatocellular adenoma; MRI, magnetic resonance imaging; T2W, T2-weighted.

### ***Malignant transformation***

Malignant transformation of HCA to HCC is rarely reported, but is an accepted risk, particularly when the diameter of the adenoma exceeds 5 cm (Figure 2) [Stoot et al. 2010; Grazioli et al. 2013]. In a systematic review, Stoot and colleagues reported an overall frequency of malignant transformation of 4.2% for HCAs [Stoot et al. 2010] (67 of 1635 HCAs, CI 0–100%). Only three cases showed malignant transformation for tumors <5 cm in diameter, which represented 4.4% of the total number of HCCs arising from HCAs (3 out of 67). As suggested for the internal bleeding data, these reports should be interpreted with caution.



**Figure 2.** A 32-year-old female using oral contraceptives with a 11 cm lesion in the liver. Based on MRI this lesion was compatible with a HCA. However, both on T2-weighting (Figure 1(a)) as on the images after contrast injection the lesion appeared heterogeneous with a focus of diffusion restriction (typically a low ADC value, (b)). Diffusion restriction is thought to be a typical sign of malignancy in liver lesions. Based on the findings above and because the lesion was larger than 5 cm, the lesion was resected 3 months later. On gross pathology there was a focal nidus (Figure 1(c), arrow, concordant with the nidus on MRI) which appeared to be an HCC in a HCA ('nodule-into-nodule'). On histology (H-E  $\times 25$ , (d)) at the interface HCA/HCC, the upper part of the tumor showed proliferation of hepatocytes without obvious cytological anomalies, intermingled with thin/ isolated vessels (down side of dotted line), favoring an HCA. 'Nodule-into-nodule' consists of clearer cells with mild atypia (above dotted-line, (d)), disorganized or decreased reticulin fibers (e) and obvious positivity for Glypican-3 (f), favoring an HCC (well differentiated). (g) GS-staining pattern at the periphery of the HCA. Glypican-3, Serum Amyloid A and CRP were negative in the HCA.  $\beta$ -catenin staining showed only membranous expression. Based on the above we interpreted this HCA as a b-HCA [van Aalten 2011b]. CRP, C-reactive protein; HCA, hepatocellular adenoma; HCC, hepatocellular carcinoma; GS staining, glutamine synthetase immunostaining indicative of b-HCA; MRI, magnetic resonance imaging.

**Table 3.** Summary of the findings of an earlier review of 12 articles in which the percentage of hemorrhaged HCAs, and minimal lesion sizes, were reported. Hemorrhage occurred mostly in larger HCAs (>5 cm; minimally 42.2%), but smaller lesions also showed some bleeding (range 8.3–11.5%).

Series	Patients with hemorrhaged HCA	Size of smallest HCA (cm)	Percentage <5 cm of total (%)
[Reddy et al. 2001]	3 of 25	4	–
[Hung et al. 2001]	4 of 25	4.2	–
[Toso et al. 2005]	10 of 25	1.7	–
[Cho et al. 2008]	12 of 41	1	8.3 (1/12)
[Bioulac-Sage et al. 2009]	23 of 128	<5	–
[Edmondson et al. 1976; Dokmak et al. 2009]	26 of 122	<5	11.5 (3/26)
[Edmondson et al. 1976]	10 of 42	>5	0
[Leese et al. 1988]	2 of 24	5	0
[Ault et al. 1996]	4 of 12	6	0
[Closset et al. 2000]	7 of 16	7	0
[Deneve et al. 2009]	31 of 124	>5	0
[Chung et al. 1995]	–	5	0

HCA, hepatocellular adenoma.

Of the four HCA subtypes, (exon 3) b-HCAs are known to trigger a potent mitogenic signaling pathway that is prominent in HCC [Zucman- Rossi et al. 2007; Chu and Moon, 2013; Pilati et al. 2014], which suggests a positive correlate between the two. Zucman-Rossi and colleagues reported an incidence of HCCs, or borderline malignant tumors in b-HCAs, of up to 46%; this malignant progression was seldom seen in other subtypes [Zucman-Rossi et al. 2007], and was over-represented for male patients (5 cases, 38%;  $p = 0.02$ ) [Hussain et al. 2006]. Since the  $\beta$ -catenin pathway can also be activated in IHCA [van Aalten et al. 2011], both the b-HCA and IHCA subtypes may necessitate more aggressive treatment than either the H-HCA or UHCA, although the clinical relevance of this determination has yet to be broadly accepted. In follow up, malignant progression of HCA to HCC has only rarely been demonstrated, with questionable quality of the imaging data for those rare, reported cases. Therefore it is presently difficult to prove that HCC is a transition from HCA, although the presence of  $\beta$ -catenin has been suggested as a criterion for the selection of HCA, or well-defined HCC, for resection [Bioulac-Sage et al. 2013]. Interestingly, Figure 2 shows a lesion with a typical nodule-in-nodule appearance which suggests a form of transition from HCA to HCC. Another problem is that corroboration of the pathology is seldom available, due to the fact that biopsies of HCA are rarely performed, with diagnoses generally made with MRI [Hussain et al. 2006]. A final diagnosis of b-HCA based solely on MRI findings would be helpful, but the MRI findings published to date for this subtype are sparse (Table 2). Finally, it should be mentioned that HCA shows a higher risk of malignant transformation in men [Farges et al. 2011]. In these cases, the possibility of hepatitis, an underlying glycogen storage disease (Figure 3), or sex



steroid hormone abuse, should all be considered as all predispose to HCC [Yoneda et al. 2012]. A more aggressive treatment is advised in such cases, even for lesions <5 cm.

Finally, according to the recent literature, H-HCA almost never degenerate into HCC, although some very rare cases have been reported [Stueck et al. 2015]. [Semin Liver Dis. 2015 Nov;35(4):444-9. Hepatocellular Carcinoma Arising in an HNF-1 $\alpha$ -Mutated Adenoma in a 23-Year-Old Woman with Maturity-Onset Diabetes of the Young: A Case Report. Stueck AE, Qu Z, Huang MA, Campreciós G, Ferrell LD, Thung SN.] The low risk of H-HCA degeneration may help to simplify the management of liver adenomas as will be discussed later.

As for clinical application, mainly b-HCA and IHCA are prone to malignant degeneration, and mostly if >5 cm. In these instances, invasive treatment is recommended.

### *Pregnancy*

Women with HCA who are pregnant, or wish to become pregnant (Figure 4), should be closely monitored for HCA size (with ultrasound or MRI) during their pregnancy, due to the tendency of the lesion to grow, especially during the third trimester when high levels of estrogens are present [Cobey and Salem, 2004]. Hormone-induced growth, and possible rupture, may result in potentially lethal complications for the mother and unborn child. Treatment of HCA during pregnancy may be indicated when the lesion shows signs of growth or bleeding, however specific figures for the risk of HCA complication during pregnancy are not yet available.

Whether some subtypes are more prone to complications during pregnancy is not known, mainly because the majority is diagnosed non-invasively.

The choice of follow up, surgery, radiofrequency ablation (RFA) or transcatheter arterial embolization (TAE) for the treatment of HCAs in pregnancy is often a matter of debate. Surgery of lesions located at the periphery of the liver can be performed safely within the first or second trimester, and will usually be indicated by the size of the lesion, and its change in size. Radiation exposure or exposure to iodinated contrast media during RFA or TAE may be contraindicated during the early phase of pregnancy, with the treatment of smaller lesions not being indicated. Given the increased risk of hemorrhage in larger HCAs (>5 cm), or when a previous pregnancy was complicated by either minor or major bleeding, we currently advocate a preemptive treatment strategy before pregnancy, as proposed by van Aalten and colleagues [Broker Broker and colleagues 2012]. Whenever a HCA is discovered during pregnancy, the second trimester is the optimal moment for invasive treatment, if indicated, as anesthesia is well tolerated at this stage, and the fetus is not yet so large as to interfere with liver surgery [Parangi et al. 2007].



### ***Liver adenomatosis***

Hepatic adenomatosis (HCAs more than 10) is regarded by some authors as a different entity [Barthelmes and Tait, 2005; Frulio et al. 2014]. There seems not to be a strong association with estrogen or anabolic steroid use [Chiche et al. 2000; Grazioli et al. 2000]. However, there is a strong association with glycogen storage disease [Chiche et al. 2000; Frulio et al. 2014]. Mostly, these adenomas are of the H-HCA and IHCA subtypes [Frulio et al. 2014]. The nodules in hepatocellular adenomatosis are often of the same subtypes. Although one might assume that multiple HCAs increase the propensity for lesional bleeding, previous data have shown no significant difference in macroscopic bleeding between single and multiple HCAs ( $p < 0.155$ ) [Dokmak et al. 2009]. According to literature there seems to be no indication to suggest that the risk of malignant transformation is increased in hepatic adenomatosis compared with solitary HCAs [Barthelmes and Tait, 2005]. However, hepatic adenomatosis are more often found in glycogen storage disease and in man [Chiche et al. 2000], and as such at risk for increased malignant potential. Presently, there is no systematic review available which evaluates the malignant potential of hepatic adenomatosis. As for clinical management, and there are no data suggesting that hepatic adenomatosis should be treated differently from solitary HCAs.

### **Biopsy in the management of HCA**

Since the introduction of the HCA subclassification system, several authors have attempted to further refine the diagnostic work-up using additional techniques, including immunohistochemistry [Zucman-Rossi et al. 2006]. The primary motivation for the introduction of additional biopsies was the prospect of identifying HCAs with greater malignant potential (such as exon-3 $\beta$ -catenin mutated HCAs). There seem to be no unique MRI features with which to assign a b-HCA subtype risk, which offers one argument for the expansion of the use of diagnostic biopsy in order to arrive at a correct diagnosis.

However, at present there is no consensus regarding the diagnostic work-up of HCA [Nault et al. 2013; Marrero et al. 2014]. Nault and colleagues regard histologic analysis as the backbone of HCA diagnosis, with the detection or exclusion of b-HCA as the main input [Nault et al. 2013]. They argue that biopsy should be offered in all cases of HCA <5 cm with no typical MRI sign of H-HCA. Lesions >5 cm do not require biopsy since they are all preferably resected. In our opinion, and in accordance with recent American College of Gastroenterology guidelines for liver lesions, the diagnostic workup of suspected HCA should be based primarily on MRI findings, with biopsy in cases where the lesion cannot be clearly differentiated from FNH [Marrero et al. 2014]. Other indications for biopsy are an atypical presentation of the HCA on imaging, with the main differential diagnosis being HCC in a noncirrhotic liver [Marrero et al. 2014].

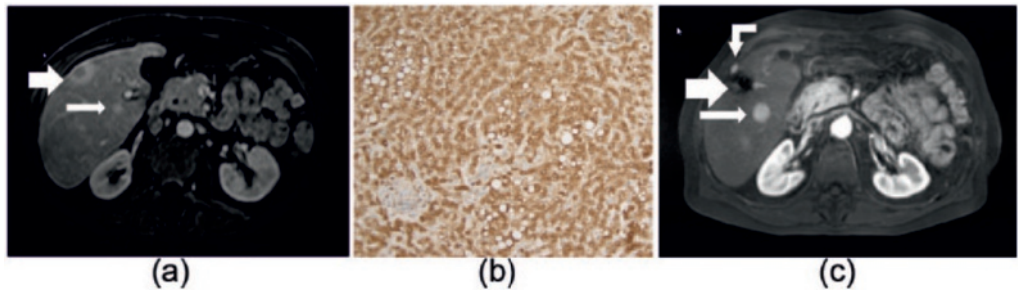
The biopsy of all HCAs (with the exclusion of typical H-HCAs based on MRI) found by imaging would be impractical. Most patients with HCA are young, with minor symptoms on malignant progression; invasive procedures should preferably be avoided. While the risk of bleeding

complications is very low when using an 18G core needle biopsy (0.6%) [Haage et al. 1999; Kadri Aribas et al. 2010; Aribas et al., 2012], the risks are not negligible, and deaths due to bleeding complications have been reported [Stattaus et al. 2007]. In our practice a biopsy has not been performed to date, except where the diagnosis of a specific adenoma subtype was expected to alter clinical management.

Unquestionably, a biopsy for further characterization may add important information in welldefined cases. For example, a biopsy with the additional help of immunostaining could facilitate better discrimination between HCA and FNH, as shown in a large retrospective study in France where the investigators compared biopsies against a final diagnosis based on surgical resection [Bioulac-Sage et al. 2012]. A total of 239 needle biopsies were compared with the final diagnosis made on resection. A difference in sensitivity of 74.3% with immunostaining versus 58.6% achieved with routine analyses without GS or other molecular features was reported [Bioulac- Sage et al. 2012]. These data suggest that immunostaining should be made available in centers that routinely treat HCAs.

What is the best approach in cases where differentiation between HCA and HCC is not evident based on MRI? In cases where there is a major suspicion of malignancy (e.g. HCC in noncirrhotic liver) based on a combination of clinical findings, size of the lesion, increased serum  $\alpha$ -fetoprotein, and MRI findings (such as heterogeneous presentation with heterogeneous enhancement, washout, and true capsule) (Figure 2), resection without prior biopsy can be recommended. Although biopsy of each suspect lesion would undoubtedly help in detecting HCC, this approach may be impractical due to the significant level of false-negative findings, the chances of seeding (Figure 3), and the enhanced risk of bleeding, which is particularly relevant when multiple biopsies are taken. Furthermore, in cases with a typical presentation, a biopsy will not influence the decision to remove the lesion. Therefore, we suggest to biopsy in selected cases only. Interestingly, the HCC literature documents a similar debate on whether it is acceptable to diagnose HCC in a cirrhotic liver based solely on MRI findings, or whether the use of routine biopsies should be advocated for all suspected lesions in patients with liver cirrhosis [Sherman and Bruix, 2015; Torbenson and Schirmacher, 2015]. Even in high grade dysplastic nodes, follow up by imaging is still preferred above biopsies.

As for daily practice, we recommend biopsy only in very selected cases where HCA cannot be differentiated from FNH with any imaging modality. The clinical repercussion of a wrong diagnosis of either HCA or FNH can have a large influence on a patient's future in terms of treatment and prognosis. When there are signs of malignancy patients should preferentially be forwarded to an experienced referral center for further evaluation. One should be aware that in some cases MRI or biopsy will be unable to differentiate between HCA and well-differentiated HCA.



**Figure 3.** A 50-year-old male with multiple hypervascular lesions. These lesions were diagnosed as HCA or HCC based on imaging and clinical (glycogen storage disease) findings. (a) An axial MR image, with T1 weighting, after contrast injection in the arterial phase. In segment 5, a small lesion with a cystic central portion (large arrow) was biopsied, and subsequently diagnosed as HCC following positive GS staining with negative  $\beta$ -catenin staining. Posteriorly, a second, smaller lesion was visualized (<1 cm, small arrow). Histologic sample of a lesion with diffuse GS-positivity (b). Axial MR image with T1-weighting after contrast injection in the arterial phase (c). In this image, taken 3 years later, the second lesion has grown (now 3 cm, small arrow). The large arrow shows the resection/ablated part of the liver (from lesion 1). A new hypervascular lesion (curved arrow) was also detected outside the liver, which proved to be a trajectory metastasis. These lesions (large arrow, curved arrow) were successfully ablated. This patient is currently being followed at regular, short intervals, and is on the waiting list for a liver transplantation. GS staining, glutamine synthetase immunostaining indicative of b-HCA, even with a negative  $\beta$ -catenin staining; HCA, hepatocellular adenoma; HCC, hepatocellular carcinoma; MR, magnetic resonance.

## Treatment options for HCA

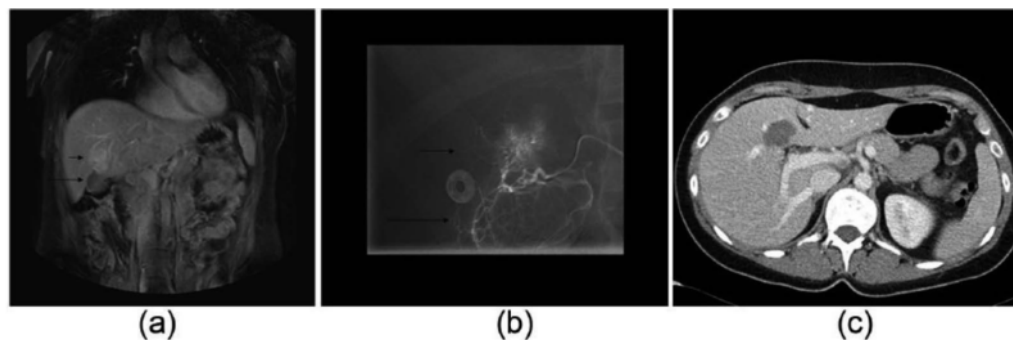
Historically, HCAs were treated with a wait-and-see policy, with surgical intervention preferred for larger (>5 cm) growing tumors. Current management options for HCAs may also include RFA, and TAE, mainly due to the advantages of these minimally-invasive techniques. In the following section we will discuss routine, as well as less commonly used HCA treatment options, with a focus on minimally-invasive, image-guided, treatment options.

### Conservative treatment

When HCAs are <5 cm, or regress (to <5 cm) following cessation of OCs, with no further growth detected, a wait-and-see policy is warranted. Although no widely accepted approach has yet been published, we prefer to schedule a patient for follow up, including MRI, or ultrasound in a yearly follow up until menopause.

### Surgery

Surgery has long been considered the treatment of choice because complete surgical removal of the lesion can be achieved in a controlled and relatively safe manner. Elective surgical resection is considered for all lesions >5 cm in diameter.



**Figure 4.** A 25-year-old female with a MRI diagnosis of single HCA, probably inflammatory subtype. As the patient wished to become pregnant, despite growth of her HCA, a decision to treat with TAE was taken. Coronal MR image with T1-weighting of the upper abdomen (a). A hypervascular HCA is indicated (small arrow), adjacent to the gallbladder (long arrow). Ablation was contraindicated due to the close proximity of the gallbladder. Angiogram (b) before TAE showing an arterial tumor 'blush' in the HCA (short arrow), with the gallbladder perfused by the same local hepatic artery division (long arrow). This finding contraindicated TAE due to the possibility of gallbladder necrosis following infarction. Instead, a decision to operate was made, with resection of the gallbladder, and subsequent intraoperative RFA of the HCA. Axial postoperative CT image after contrast injection in the venous phase (c). The gallbladder was resected, in combination with intraoperative RFA (arrow). CT, computerized tomography; HCA, hepatocellular adenoma; RFA, radiofrequency ablation; TAE, transcatheter arterial embolization.

With a mortality of 1.1% (review by Lin and colleagues,  $n = 170$ ), surgery is a relatively safe procedure. In one review, 93% of patients with ruptured, or nonruptured HCAs, were primarily treated with surgery, with complications that included two deaths, one biloma, one bile leakage, one infection, and one case of sepsis [Lin et al. 2011]. In another, single-center retrospective analysis of 41 cases, no perioperative mortality was found, and only minor complications arose. These included pleural effusion requiring drainage ( $n = 2$ ), pneumonia ( $n = 1$ ), and wound infection ( $n = 1$ ) [Cho et al. 2008].

In the latter study, nine cases were operated on laparoscopically, a technique that is increasingly popular, where appropriate. De'Angelis and colleagues described 62 HCA patients who underwent either an open procedure or laparoscopy [De'Angelis et al. 2014]. They found no difference in postoperative morbidity and zero mortality, with no longterm complications or recurrence of HCA. However, patients with smaller lesions were preferentially treated with laparoscopy (68 versus 9). These authors concluded that laparoscopic liver resections may be limited by lesion size and location, and that the technique requires advanced surgical skills. Given the precision required, robotic surgery may prove to be useful in the future, and could reduce complications; we await further evaluation of its efficacy [Jackson et al. 2015].

In rare circumstances, the treatment of HCA may also involve liver transplantation, a procedure described in a case report by Vennarecci and colleagues [Vennarecci et al. 2013]. Obesity,

steatosis, and diabetes, are frequent co-morbidities in patients with HCAs, particularly the inflammatory subtype. These factors, and especially obesity, make surgery less attractive. For those patients who are poor candidates for surgery (centrally located lesions, multiple adenomas, or morbid obesity), RFA and TAE may instead be offered.

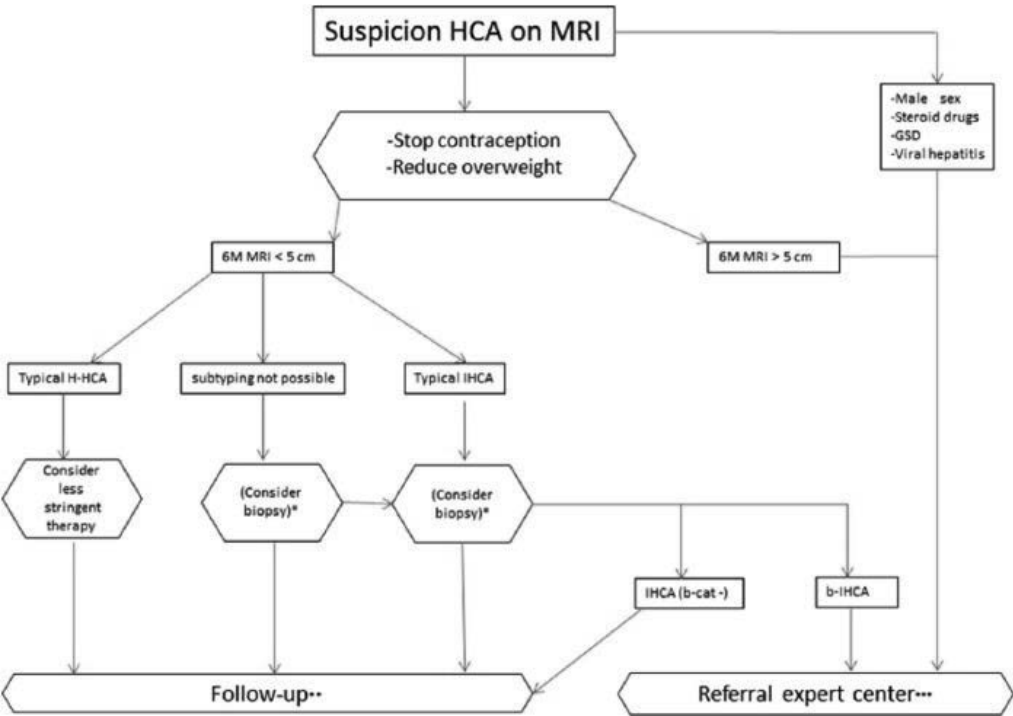
### ***Radiofrequency ablation***

RFA is a minimally invasive technique used in the treatment of HCC, other liver lesions such as colorectal metastases [Solbiati et al. 2001; Cabibbo et al. 2013], and HCAs [van Aalten et al. 2010; van Vledder et al. 2011]. Laparoscopic RFA, or perioperative RFA, may also be considered when the anatomical location [i.e. close proximity to the bowel or gallbladder (Figure 3)] leads to an increased risk using a percutaneous approach. The use of RFA in the treatment of HCAs has only been described in small case series.

Van Vledder and colleagues described one case series including 45 HCAs, in 18 patients, that were ablated in 32 RFA sessions (open,  $n = 4$ ; percutaneous,  $n = 28$ ) [Van Vledder et al. 2011]. A total of 26 of 45 HCAs were successfully treated in one RFA session, with no visible residual disease. A further nine HCAs were totally ablated following a second RFA session. There were 3 HCAs that required 3 or more RFA sessions, with all but 7 of the 45 HCAs totally ablated after 3 or more sessions. The treated HCAs had a median size of 3.0 cm (ranging from 0.8 to 7.3 cm). Only minor complications were attributable to the RFA procedure; none of which required additional intervention (class A according to the Society of Interventional Radiology scoring system for complications). A single class D major complication was reported; a cerebrovascular accident during open surgery combined with RFA. Though severe, this complication was linked to anesthesiological and hemodynamic changes during laparotomy, rather than the RFA procedure. In conclusion, RFA can be effectively and safely used in the treatment of HCA, although multiple sessions may be required for larger lesions.

In a review of HCA cases reported between 1998– 2008, Lin and colleagues identified 356 HCA patients in reports from China, Europe, North America, and South-East Asia [Lin et al. 2011]. Only 14 (3.9%) of these cases were treated with RFA, and no severe complications were reported. However, no results in terms of efficacy were provided. In 2008, Rhim and colleagues assessed the therapeutic efficacy and safety of RFA for HCAs [Rhim et al. 2008], and reported their initial experience in 10 patients with 12 HCAs. Tumor sizes ranged from 1.5 to 4.5 cm. As no complications were reported after RFA, and no progression or recurrence was noted, RFA was considered a safe and effective treatment option. A minimal ablative margin of 5 mm is recommended during the radical treatment of lesions when using thermal ablation of HCCs. It is presently unclear if a similar margin should be applied to HCAs, as these lesions are assumed to be benign. No data are yet available regarding the ideal ablative margin during thermal ablation of HCAs. In our opinion, volume reduction is more important than an ablation margin, as the former correlates strongly with a risk of bleeding, and malignant transformation.

Follow-up imaging after both RFA, and TAE, is routinely performed in our institution by MRI, 6 months after treatment.



**Figure 5.** The management decision tree used in our tertiary academic medical center. This decision tree may not be appropriate for all patients; for some, a customized approach should be considered. \*One option is to biopsy those lesions where a subtyping diagnosis by imaging is impossible to achieve, or those lesions with typical signs of IHCA. Currently, this option is considered impractical given the large number of biopsies involved.[AQ: 8] \*\*Follow up is advised initially, at 6-monthly intervals. Thereafter, if the lesion shows no further alteration, follow up can be stopped, or repeated yearly until menopause. If the lesion is a typical H-HCA, follow up can be less stringent, possibly involving sonography, or MRI without contrast. \*\*\*Referral to an expert center is advised for the evaluation of any indication requiring intervention. This decision should be taken with consideration of contraindications (obesity, diabetes, centrally located tumor, ASA classification). Treatment can be primarily surgical, and in selected cases, RFA or local embolization. b-HCA,  $\beta$ -catenin activated HCA; GSD, glycogen storage disease; HCA, hepatocellular adenoma; H-HCA, HNF1A-mutated HCA; IHCA, inflammatory HCA; M, months; RFA, radiofrequency ablation.

As HCAs requiring treatment are generally large (>5 cm), a promising alternative for RFA may be microwave ablation (MWA). Based on preliminary data, MWA was shown to produce larger ablation zones, in less time, in patients treated for HCC and colorectal metastases. MWA delivers high frequency microwaves (0.9-2.4 GHz) into tumor tissue, which causes fast

spinning of molecules and thus destroys tissue. No data are available concerning efficacy and complications after MWA. MWA has specific advantages over RFA, such as larger ablation zones, higher treatment temperatures, and less susceptibility to local cooling by adjacent large blood vessels (heatsink). Although larger zones of ablation can probably be achieved by using single-electrode needles and MWA to treat HCAs, to the best of our knowledge, no data currently exists to substantiate this idea.

### ***Transcatheter arterial embolization***

As HCAs are hypervascular arterial lesions, bleeding may be treated by selective transcatheter arterial embolization (TAE) in cases where patients present with hemodynamic instability. Embolization of HCAs is a safe but relatively challenging procedure due to multiple small feeding vessels [van Aalten et al. 2010]. Nonetheless, in cases of spontaneous rupture and bleeding, TAE should be considered as a first line treatment as it is highly successful and minimally invasive in an acute setting. Although high success rates have been described for TAE, there is only a sparse literature comparing TAE with either surgery, or conservative management. In one study Karkar and colleagues described 52 patients with 100 HCAs, of which 37% were treated with TAE [Karkar et al. 2013]. In most of these cases TAE was performed in a (semi) elective setting, with rupture and hemorrhage as indications in 20%, and suspected malignancy in 56%. Success rates of up to 92% were claimed for TAE (32), and of the 37 HCAs embolized, only 3 required secondary interventions (8.1%). All other embolized lesions were treated successfully; some disappeared (5/34), most decreased in size (22/34), or remained stable (7/34). Recurrence rates were also low. It is worth noting that the HCAs embolized were relatively small, with a median diameter of 2.6 cm. However, we feel that resection is indicated if malignancy is suspected and no contraindications for surgery exist. In a report by Erdogan and colleagues, six HCAs were primarily embolized [Erdogan et al. 2007], four because of bleeding, and two electively, 1 year after bleeding. No complications were reported, and all HCAs ceased bleeding. A total of two of the lesions were resected after embolization, two regressed visibly on follow-up imaging, and two HCAs were only seen after resorption of hematoma on follow-up imaging. These last two patients were managed with a wait-and-see policy. In a retrospective study by Dheodhar and colleagues, 17 embolizations were successfully performed in eight patients [Deodhar et al. 2011], with five patients undergoing more than one embolization. The mean size of the treated HCAs was 3.6 cm, and regression was noted in all embolized HCAs after embolization. As noted by Karkar and colleagues, TAE may also be used in an elective setting where no acute intervention is needed [Karkar et al. 2013]. This approach is of clinical interest and deserves further consideration.

### **Proposed management strategy**

One of the major discussions on the management strategy of hepatocellular adenomas involves the clinical application of these recent findings in the dynamic field of adenoma subtyping. How should we take into account these new insights into daily practice? In our view, more data are needed to implement this subclassification in the diagnosis and treatment of

adenomas, balancing the risk of an invasive liver biopsy with the additional benefits in terms of individualized therapy and prognostic stratification. A major effort should be made by expert centers involved in the diagnosis and treatment of hepatocellular adenomas to work on this collaboratively, preferably in research setting, to gather more data on the potential benefits for an individual patient.

Based on our review of the current literature, we propose a management strategy applicable to most cases in which there is a suspicion of HCA (Figure 5). This decision tree may not be appropriate for all patients; for some, a more customized approach may be required. In standard situations, mainly when a lesion is  $>5$  cm, OCs should be stopped and MRI performed after 6 months. If the lesion has contracted to  $<5$  cm, clear signs of an H-HCA should be ruled in or out (Figure 1). In scenarios where H-HCAs are subsequently identified, therapy can be less aggressive as inherent malignant progression is very low. Follow up is then advised, initially every 6 months, and if the lesion shows no further alteration, follow up can be stopped or simply repeated yearly until menopause [Marrero et al. 2014]. Since typical H-HCAs are easily identified using out-of-phase MRI sequences, intravenous contrast can be obviated at follow up. A second option is to apply sonography during follow up which is cheaper and less inconvenient for patients. For small lesions ( $<5$  cm) that are categorized as IHCA, therapy should ordinarily not be altered (in standard cases). However, some may opt for a biopsy in order to exclude  $\beta$ -catenin mutation. This could also be the case if a subclassification cannot be made with MRI. For larger lesions ( $>5$  cm) with a  $\beta$ -catenin mutation or if the patient has an aggravating status such as male sex, steroid use, glycogen storage disease, or underlying viral hepatitis, intervention may be the first alternative. Treatment can be primarily surgical, and, in selected cases, RFA or TAE may be used. Depending on the underlying risks (obesity, diabetes, centrally located tumor), the best option would be to evaluate these patients in an expert referral center.

## Conclusion

MRI is the preferred tool in the management of HCA, its current principle use being size evaluation (cutoff 5 cm), identification of signs of malignancy, and exclusion of H-HCAs, recognized for their benign course and permitting a conservative approach. Until now, there is no reliable MRI characteristic to diagnose non-invasively b-HCA, being the most important lesion to diagnose as it may have the highest malignant potential.

Conservative management remains the strategy of choice for uncomplicated small HCAs, and surgery may be indicated if imaging shows heterogeneous signal and growing smaller lesions suspected of being highly-differentiated HCC. Further prospective cohort studies are warranted to support the choices made between these treatment strategies and to determine the role of biopsy in the subclassification of HCAs. In cases where a HCA requires treatment, and surgical resection of smaller lesions ( $<3$  cm) carries an unacceptable risk, RFA or TAE may be considered.



## **Funding**

This research received no specific grant from any funding agency in the public, commercial, or not for-profit sectors.

## **Conflict of interest statement**

The authors declare that there is no conflict of interest.

## References

- Agarwal, S., Fuentes-Orrego, J., Arnason, T., Misdraji, J., Jhaveri, K., Harisinghani, M. et al. (2014) Inflammatory hepatocellular adenomas can mimic focal nodular hyperplasia on gadoxetic acid-enhanced MRI. *AJR Am J Roentgenol* 203: W408-414.
- Aribas, B., Arda, K., Ciledag, N., Aktas, E., Yakut, F., Kavak, S. et al. (2012) Accuracy and safety of percutaneous US-guided needle biopsies in specific focal liver lesions: comparison of large and small needles in 1300 patients. *Panminerva Med* 54: 233-239.
- Ault, G., Wren, S., Ralls, P., Reynolds, T. and Stain, S. (1996) Selective management of hepatic adenomas. *Am Surg* 62: 825-829.
- Barthelmes, L. and Tait, I. (2005) Liver cell adenoma and liver cell adenomatosis. *HPB (Oxford)* 7: 186-196.
- Baum, J., Bookstein, J., Holtz, F. and Klein, E. (1973) Possible Association between benign hepatomas and oral contraceptives. *Lancet* 2: 926-929.
- Bieze, M., Phoa, S., Verheij, J., van Lienden, K. and van Gulik, T. (2014) Risk factors for bleeding in hepatocellular adenoma. *Br J Surg* 101: 847-855.
- Bieze, M., van Den Esschert, J., Nio, C., Verheij, J., Reitsma, J., Terpstra, V. et al. (2012) Diagnostic accuracy of mri in differentiating hepatocellular adenoma from focal nodular hyperplasia: prospective study of the additional value of gadoxetate disodium. *AJR Am J Roentgenol* 199: 26-34.
- Bioulac-Sage, P., Balabaud, C., Bedossa, P., Scoazec, J., Chiche, L., Dhillon, A. et al. (2007a) Pathological diagnosis of liver cell adenoma and focal nodular hyperplasia: Bordeaux update. *J Hepatol* 46: 521-527.
- Bioulac-Sage, P., Balabaud, C. and Zucman-Rossi, J. (2010) Focal nodular hyperplasia, hepatocellular adenomas: past, present, future. *Gastroenterol Clin Biol* 34: 355-358.
- Bioulac-Sage, P., Cubel, G., Taouji, S., Scoazec, J., Leteurtre, E., Paradis, V. et al. (2012) Immunohistochemical markers on needle biopsies are helpful for the diagnosis of focal nodular hyperplasia and hepatocellular adenoma subtypes. *Am J Surg Pathol* 36: 1691-1699.
- Bioulac-Sage, P., Laumonier, H., Couchy, G., Le Bail, B., Sa Cunha, A., Rullier, A. et al. (2009) Hepatocellular adenoma management and phenotypic classification: the Bordeaux experience. *Hepatology* 50: 481-489.
- Bioulac-Sage, P., Rebouissou, S., Thomas, C., Blanc, J., Saric, J., Sa Cunha, A. et al. (2007b) Hepatocellular adenoma subtype classification using molecular markers and immunohistochemistry. *Hepatology* 46: 740-748.
- Bioulac-Sage, P., Sempoux, C., Possenti, L., Frulio, N., Laumonier, H., Laurent, C. et al. (2013) Pathological diagnosis of hepatocellular cellular adenoma according to the clinical context. *Int J Hepatol* 2013: 253261.
- Blanc, J., Frulio, N., Chiche, L., Sempoux, C., Annet, L., Hubert, C. et al. (2015) Hepatocellular adenoma management: call for shared guidelines and multidisciplinary approach. *Clin Res Hepatol Gastroenterol* 39: 180-187.
- Broker, M., IJzermans, J., van Aalten, S., De Man, R. and Terkivatan, T. (2012) The management of pregnancy in women with hepatocellular adenoma: a plea for an individualized approach. *Int J Hepatol* 2012: 725735.
- Cabibbo, G., Maida, M., Genco, C., Alessi, N., Peralta, M., Butera, G. et al. (2013) Survival of patients with hepatocellular carcinoma (HCC) treated by percutaneous radio-frequency ablation (RFA) is affected by complete radiological response. *PLoS One* 8: e70016.
- Chiche, L., Dao, T., Salame, E., Galais, M., Bouvard, N., Schmutz, G. et al. (2000) Liver adenomatosis: reappraisal, diagnosis, and surgical management: eight new cases and review of the literature. *Ann Surg* 231: 74-81.
- Cho, S., Marsh, J., Steel, J., Holloway, S., Heckman, J., Ochoa, E. et al. (2008) Surgical management of hepatocellular adenoma: take it or leave it? *Ann Surg Oncol* 15: 2795-2803.
- Chu, H. and Moon, W. (2013) Beta-catenin activated hepatocellular adenoma. *Clin Mol Hepatol* 19: 185-189.

- Chung, K., Mayo-Smith, W., Saini, S., Rahmouni, A., Golli, M. and Mathieu, D. (1995) Hepatocellular adenoma: MR imaging features with pathologic correlation. *AJR Am J Roentgenol* 165: 303-308.
- Closset, J., Veys, I., Peny, M., Braude, P., van Gansbeke, D., Lambilliotte, J. et al. (2000) Retrospective analysis of 29 patients surgically treated for hepatocellular adenoma or focal nodular hyperplasia. *Hepatogastroenterology* 47: 1382-1384.
- Cobey, F. and Salem, R. (2004) A review of liver masses in pregnancy and a proposed algorithm for their diagnosis and management. *Am J Surg* 187: 181-191.
- De'Angelis, N., Memeo, R., Calderaro, J., Felli, E., Salloum, C., Compagnon, P. et al. (2014) Open and laparoscopic resection of hepatocellular adenoma: trends over 23 years at a specialist hepatobiliary unit. *HPB (Oxford)* 16: 783-788.
- Deneve, J., Pawlik, T., Cunningham, S., Clary, B., Reddy, S., Scoggins, C. et al. (2009) Liver cell adenoma: a multicenter analysis of risk factors for rupture and malignancy. *Ann Surg Oncol* 16: 640-648.
- Deodhar, A., Brody, L., Covey, A., Brown, K. and Getrajdman, G. (2011) Bland embolization in the treatment of hepatic adenomas: preliminary experience. *J Vasc Interv Radiol* 22: 795-799, quiz 800.
- Dokmak, S., Paradis, V., Vilgrain, V., Sauvanet, A., Farges, O., Valla, D. et al. (2009) A single-center surgical experience of 122 patients with single and multiple hepatocellular adenomas. *Gastroenterology* 137: 1698-1705.
- Edmondson, H., Henderson, B. and Benton, B. (1976) Liver-cell adenomas associated with use of oral contraceptives. *N Engl J Med* 294: 470-472.
- Erdogan, D., van Delden, O., Busch, O., Gouma, D. and van Gulik, T. (2007) Selective transcatheter arterial embolization for treatment of bleeding complications or reduction of tumor mass of hepatocellular adenomas. *Cardiovasc Intervent Radiol* 30: 1252-1258.
- Farges, O., Ferreira, N., Dokmak, S., Belghiti, J., Bedossa, P. and Paradis, V. (2011) Changing trends in malignant transformation of hepatocellular adenoma. *Gut* 60: 85-89.
- Frulio, N., Chiche, L., Bioulac-Sage, P. and Balabaud, C. (2014) Hepatocellular adenomatosis: what should the term stand for? *Clin Res Hepatol Gastroenterol* 38: 132-136.
- Grazioli, L., Federle, M., Ichikawa, T., Balzano, E., Nalesnik, M. and Madariaga, J. (2000) Liver adenomatosis: clinical, histopathologic, and imaging findings in 15 patients. *Radiology* 216: 395-402.
- Grazioli, L., Olivetti, L., Mazza, G. and Bondioni, M. (2013) MR imaging of hepatocellular adenomas and differential diagnosis dilemma. *Int J Hepatol* 2013: 374170.
- Haage, P., Piroth, W., Staatz, G., Adam, G. and Gunther, R. (1999) [CT-guided percutaneous biopsies for the classification of focal liver lesions: a comparison between 14 g and 18 g puncture biopsy needles]. *CT-Gesteuerte Perkutane Biopsien zur Klassifizierung von Fokalen Leberlasionen: Vergleich Zwischen 14 G- und 18 G-Stanzbiopsienadeln. Rofo* 171: 44-48.
- Hung, C., Changchien, C., Lu, S., Eng, H., Wang, J., Lee, C. et al. (2001) Sonographic features of hepatic adenomas with pathologic correlation. *Abdom Imaging* 26: 500-506.
- Hussain, S., van Den Bos, I., Dwarkasing, R., Kuiper, J. and den Hollander, J. (2006) Hepatocellular adenoma: findings at state-of-the-art magnetic resonance imaging, ultrasound, computed tomography and pathologic analysis. *Eur Radiol* 16: 1873-1886.
- Jackson, N., Hauch, A., Hu, T., Buell, J., Slakey, D. and Kandil, E. (2015) The safety and efficacy of approaches to liver resection: a meta-analysis. *JSLs* 19: JSLs 2015 Jan-March; 19(1):e2014.00186.
- Kadri Aribas, B., Dingil, G., Sahin, G., Nil Unlu, D., Dogan, K. et al. (2010) Accuracy and safety of percutaneous us-guided needle biopsies in liver metastasis and hemangiomas. *Minerva Gastroenterol Dietol* 56: 377-382.
- Karkar, A., Tang, L., Kashikar, N., Gonen, M., Solomon, S., Dematteo, R. et al. (2013) Management of Hepatocellular adenoma: comparison of resection, embolization and observation. *HPB (Oxford)* 15: 235-243.
- Laumonier, H., Bioulac-Sage, P., Laurent, C., Zucman-Rossi, J., Balabaud, C. and Trillaud, H. (2008) Hepatocellular adenomas: magnetic resonance imaging features as a function of molecular pathological classification. *Hepatology* 48: 808-818.

- Leese, T., Farges, O. and Bismuth, H. (1988) Liver cell adenomas. a 12-year surgical experience from a specialist hepato-biliary unit. *Ann Surg* 208: 558-564.
- Lin, H., van Den Esschert, J., Liu, C. and van Gulik, T. (2011) Systematic review of hepatocellular adenoma in china and other regions. *J Gastroenterol Hepatol* 26: 28-35.
- Marrero, J., Ahn, J. and Rajender Reddy, K. and American College of Gastroenterology (2014) ACG clinical guideline: the diagnosis and management of focal liver lesions. *Am J Gastroenterol* 109: 1328-1347; quiz 1348.
- McInnes, M., Hibbert, R., Inacio, J. and Schieda, N. (2015) Focal nodular hyperplasia and hepatocellular adenoma: accuracy of gadoxetic acid-enhanced MR imaging: a systematic review. *Radiology* 277: 413-423.
- Nault, J., Bioulac-Sage, P. and Zucman-Rossi, J. (2013) Hepatocellular benign tumors-from molecular classification to personalized clinical care. *Gastroenterology* 144: 888-902.
- Paradis, V., Champault, A., Ronot, M., Deschamps, L., Valla, D., Vidaud, D. et al. (2007) Telangiectatic adenoma: an entity associated with increased body mass index and inflammation. *Hepatology* 46: 140- 146.
- Parangi, S., Levine, D., Henry, A., Isakovitch, N. and Pories, S. (2007) Surgical Gastrointestinal Disorders During Pregnancy. *Am J Surg* 193: 223-232.
- Pilati, C., Letouze, E., Nault, J., Imbeaud, S., Boulai, A., Calderaro, J. et al. (2014) Genomic profiling of hepatocellular adenomas reveals recurrent FRKactivating mutations and the mechanisms of malignant transformation. *Cancer Cell* 25: 428-441.
- Reddy, K., Kligerman, S., Levi, J., Livingstone, A., Molina, E., Franceschi, D. et al. (2001) Benign and solid tumors of the liver: relationship to sex, age, size of tumors, and outcome. *Am Surg* 67: 173-178.
- Rhim, H., Lim, H., Kim, Y. and Choi, D. (2008) Percutaneous radiofrequency ablation of hepatocellular adenoma: initial experience in 10 patients. *J Gastroenterol Hepatol* 23: e422-e427.
- Ronot, M., Bahrami, S., Calderaro, J., Valla, D., Bedossa, P., Belghiti, J. et al. (2011) Hepatocellular adenomas: accuracy of magnetic resonance imaging and liver biopsy in subtype classification. *Hepatology* 53: 1182-1191.
- Sherman, M. and Bruix, J. (2015) Biopsy for liver cancer: how to balance research needs with evidencebased clinical practice. *Hepatology* 61: 433-436.
- Solbiati, L., Livraghi, T., Goldberg, S., Ierace, T., Meloni, F., Dellanoce, M. et al. (2001) Percutaneous radio-frequency ablation of hepatic metastases from colorectal cancer: long-term results in 117 patients. *Radiology* 221: 159-166.
- Stattaus, J., Kuhl, H., Hauth, E., Kalkmann, J., Baba, H. and Forsting, M. (2007) [Liver biopsy under guidance of multi slice computed tomography: comparison of 16g and 18g biopsy needles]. *Leberbiopsie Mit Hilfe der Mehrschichtcomputertomographie: Vergleich Zwischen 16-G- und 18-G-Biopsienadeln. Radiologe* 47: 430-438.
- Stoot, J., Coelen, R., de Jong, M. and Dejong, C. (2010) Malignant transformation of hepatocellular adenomas into hepatocellular carcinomas: a systematic review including more than 1600 adenoma cases. *HPB (Oxford)* 12: 509-522.
- Thomeer, M., Me, E., de Lussanet, Q., Biermann, K., Dwarkasing, R., de Man, R. et al. (2014a) Genotype-phenotype correlations in hepatocellular MG Thomeer, M Broker et al. <http://tag.sagepub.com> 15 adenoma: an update of MRI findings. *Diagn Interv Radiol* 20: 193-199.
- Thomeer, M., Willemssen, F., Biermann, K., El Addouli, H., de Man, R., IJzermans, J. et al. (2014b) MRI features of inflammatory hepatocellular adenomas on hepatocyte phase imaging with liverspecific contrast agents. *J Magn Reson Imaging* 39: 1259-1264.
- Torbenson, M. and Schirmacher, P. (2015) Liver cancer biopsy: back to the future? *Hepatology* 61: 431-433.
- Toso, C., Majno, P., Andres, A., Rubbia-Brandt, L., Berney, T., Buhler, L. et al. (2005) Management of hepatocellular adenoma: solitary-uncomplicated, multiple and ruptured tumors. *World J Gastroenterol* 11: 5691-5695.
- Van Aalten, S., de Man, R., IJzermans, J. and Terkivatan, T. (2012) Systematic review of haemorrhage and rupture of hepatocellular adenomas. *Br J Surg* 99: 911-916.

- Van Aalten, S., Terkivatan, T., de Man, R., van Der Windt, D., Kok, N., Dwarkasing, R. et al. (2010) Diagnosis and treatment of hepatocellular adenoma in the netherlands: similarities and differences. *Dig Surg* 27: 61-67.
- Van Aalten, S., Thomeer, M., Terkivatan, T., Dwarkasing, R., Verheij, J., de Man, R. et al. (2011a) Hepatocellular adenomas: correlation of mr imaging findings with pathologic subtype classification. *Radiology* 261: 172-181.
- Van Aalten, S., Verheij, J., Terkivatan, T., Dwarkasing, R., de Man, R. and IJzermans, J. (2011b) Validation of a liver adenoma classification system in a tertiary referral centre: implications for clinical practice. *J Hepatol* 55: 120-125.
- Van Vledder, M., van Aalten, S., Terkivatan, T., de Man, R., Leertouwer, T. and IJzermans, J. (2011) Safety and efficacy of radiofrequency ablation for hepatocellular adenoma. *J Vasc Interv Radiol* 22: 787-793.
- Vennarecci, G., Santoro, R., Antonini, M., Ceribelli, C., Laurenzi, A., Moroni, E. et al. (2013) Liver transplantation for recurrent hepatic adenoma. *World J Hepatol* 5: 145-148.
- Yoneda, N., Matsui, O., Kitao, A., Kozaka, K., Gabata, T., Sasaki, M. et al. (2012) Betacatenin- activated hepatocellular adenoma showing hyperintensity on hepatobiliary-phase gadoxeticenhanced magnetic resonance imaging and overexpression of OATP8. *Jpn J Radiol* 30: 777-782.
- Zucman-Rossi, J., Benhamouche, S., Godard, C., Boyault, S., Grimber, G., Balabaud, C. et al. (2007) Differential Effects of Inactivated Axin1 and Activated Beta-Catenin Mutations in Human Hepatocellular Carcinomas. *Oncogene* 26: 774-780.
- Zucman-Rossi, J., Jeannot, E., Nhieu, J., Scoazec, J., Guettier, C., Rebouissou, S. et al. (2006) Genotypephenotype correlation in hepatocellular adenoma: new classification and relationship with HCC. *Hepatology* 43: 515-524.



# Chapter 11

Discussion

Samenvatting





# Discussion

This thesis discusses several aspects of female abdominal MRI and provides answers based on a series of investigations.

In part 1 **chapter 2**, we analysed the value of MRI as a predictor of parametrial invasion and advanced disease in cervical carcinoma, and then compared it with clinical examination. The primary requirement of these tests is to prevent underdiagnosis of parametrial invasion or advanced disease. If one of the tests produces a false negative result for parametrial invasion, the patient will be operated with a Wertheim procedure. Following this operation, and supported by histopathological evidence, this patient group will still have to undergo postoperative radiation, as this is considered standard care. However, this multimodal treatment (surgery, radiation therapy and chemotherapy) significantly increases morbidity in both the short and long-terms [1;2]. It is therefore very important to avoid this situation through use of a diagnostic tool that shows a particularly high sensitivity for invasion, which if detected would avoid futile surgery. Literature reports on this issue are extensive but are also partly contradictory, and this was reflected in our systematic review. It should be mentioned that there was only one multicentre study (originated in the United States), where they found a rather low sensitivity for MRI (53 %) [3]. However, this result did not have a significant effect on our funnel plots. Deeks funnel plots analyse the possible presence of publication bias with smaller studies, reporting overestimated effect sizes, being overrepresented. Moreover, the diagnostic odds were not asymmetrical between the different MRI data, which we suspect was caused by the fact that the number of inclusions in this multicentre study was not significantly higher than other studies, resulting in only a moderate effect of this study on the overall results.

Therefore, taking all data into account, we found a clear tendency for pooled sensitivity data for MRI to be significantly better than clinical examination. We therefore recommend MRI as opposed to clinical examination in the pretreatment evaluation of primary cervical carcinoma.

On the other hand, overdiagnosis of parametrial involvement will result in patient sent to radiotherapy instead of surgery, with known complications like bladder and bowel inflammation and peripheral oedema. This is also important, and therefore the specificity should be reasonably high. We found that the value of clinical examination and MRI in minimizing false positives is comparably high. Thus, the number of patients who will undergo radiotherapy instead of surgery will be low. However, verification bias is clearly present. Patients who undergo radiotherapy will have no reference standard, since there is no possibility of pathological confirmation.

In part 1 **chapter 3**, we prospectively analysed, in a multicentre design, the value of MRI in the response evaluation of cervical carcinoma treated with radiation therapy. Previous studies

showed no additional value of MRI in this setting. However, by adding diffusion imaging (DWI) to MRI as performed in our study, we found a significant increase in the specificity, with consequently fewer false positives. This result is due to the better differentiation with DWI between oedema due to radiation treatment and persistent tumour residue in the cervix. Using logistic regression analysis, we compared the value of MRI with DWI with clinical response evaluation after treatment. Our results showed that adding MRI with DWI to the response evaluation has a significant impact on the detection of residual tumour. Moreover, in cases where both clinical response evaluation and MRI with DWI is performed, only the latter seems to be determining. Therefore, our results perhaps indicate that MRI with DWI might replace clinical response evaluation in the future, if clinical examination is too difficult to perform in an outpatient clinic. However, further research is needed to consolidate this finding.

In part 1 **chapter 4**, we analysed whether an optimized MRI protocol could help to reinforce the clinical suspicion of endometriosis. In our study, we found that MRI is accurate in detecting patients with stage II up to IV endometriosis, whereas MRI seems unreliable in patients with minimal endometriosis (stage I). Although our study population was small, we evaluated our dataset per region in the pelvis in order to enhance significance. We can conclude that if MRI shows a lesion suspect for endometriosis this is probably true since this test achieves specificity above 90%, independently of region in the pelvis. The sensitivity for superficial lesions is only mediocre. However, when the MRI indication is detection of endometriosis, it is not important to detect all superficial lesions, as the detection of only a small number of superficial lesions is enough to diagnose the patient with endometriosis. Our data shows that detecting the disease is feasible when the presentation is more extensive. However, larger datasets are needed in the case of stages I or II.

In a large cohort of patients with a low prevalence of endometriosis (for instance a subfertility patient group), there will be many cases of stage I and II endometriosis. The use of MRI in this patient group to determine if endometriosis is the cause of subfertility will probably result in a large number of false negatives. However, due to high specificity, when MRI detects a lesion it can be stated with high confidence that it is a genuinely positive case.

In part 2 **chapter 5**, we carried out one of the first analyses in the literature of the value of MRI in the differentiating the various subtypes of HCA. A new MRI sign was proposed, which consists of the “atoll sign” typical for inflammatory HCA. This sign is currently used by various other authors, and seems to be a simple, recognizable tool for differentiation [4;5]. Furthermore, in line with two other studies, we found that diffuse fat accumulation in HCA is typical of L-FABP-negative HCA [6;7]. We also proposed closer follow-up if the lesion shows MRI signs of b-catenin HCA, which consists of a vaguely demarcated scar. Since that report, another study also found that this sign is indicative for this uncommon subtype of HCA [8]. Further inclusion of this subtype is still warranted to prove the value of this MRI sign.

In part 2 **chapter 6**, we showed that liver-specific contrast agents cannot always accurately differentiate HCA from FNH. Inflammatory HCA in particular can appear similar to FNH on hepatobiliary phase imaging. This was a new finding that contradicted earlier reports that a contrast agent, gadobenate dimeglumine, has very high accuracy in the differentiation of HCAs and FNHs [9]. The most likely reason for this disparity with that earlier report is that we used a more rigorous reference standard, molecular differentiation on the basis of histological specimens, which has now been adopted in standard use. The study by Grazioli et al. is therefore unreliable in the light of the current classification [9]. Furthermore, in that study inflammatory HCAs may have been misdiagnosed as telangiectatic FNH, a subtype of FNH and a term no longer in use [10].

In part 2 **chapter 7**, we reviewed the different appearances of HCAs based on typical MRI findings, the new subclassification and the use of different contrast agents. Our most important conclusion was that, in cases where the final diagnosis with hepatobiliary contrast agents (in the hepatobiliary phase) is uncertain, other typical MRI signs should be used to obtain an overall picture of the lesion.

In part 2 **chapter 8**, we discussed a recently published study, where the authors used a quantitative method to differentiate HCA from FNH with gadobenate dimeglumine. In our earlier study we demonstrated that the qualitative evaluation of the hepatobiliary phase after gadobenate dimeglumine is often incorrect. Roux et al therefore proposed a quantitative method to overcome this problem [11]. We tried to externally validate the latter results in our study in **chapter 8** but could not reproduce their findings.

In part 2 **chapter 9**, we examined whether this pitfall of hepatobiliary contrast agents is valid for other hepatobiliary contrast agents. Nowadays, only gadobenate dimeglumine and gadoxetate disodium are commercially available and they both claim to be valuable in differentiating HCA from FNH. We have already shown that the first agent is not always reliable, but in order to determine whether these contrast agents differ in validity we reviewed all patients in our hospital that underwent MRI with both contrast agents at different time points. Although our reference standard was not always optimal, clear differences became obvious. In a large number of patients the contrast agents gave different results in the same patient, with the use of one contrast agent arguing for a diagnosis of HCA, while the other contrast agent indicated the opposite. This clearly shows that one or both contrast agents are unreliable and should not be used in clinical practice. Using our reference standard (admittedly suboptimal but the best available at the time), most misdiagnoses were attributable to gadobenate dimeglumine, and in several cases this was possibly due to intrinsic hyperintensity on the precontrast scan. This finding could have implications for clinical practice. In the recent publication of clinical guidelines, the EASL practice guideline and ESGAR guideline, no preference was indicated for either contrast agent [5;12]. Based on current evidence, our view is that gadobenate should not be preferentially used for the differentiation of HCA from FNH. In our department, we

have now changed from gadobenate dimeglumine to gadoxetate disodium whenever the aim is differentiation of these entities (HCA and FNH). Within this application we know that the hepatobiliary phase is very reliable but that the dynamic phase is less sound, in contrast to gadobenate dimeglumine. Therefore, we actually prefer to use a non-specific contrast agent (with good dynamic phase) when the clinical request is broad. Only when the question of whether the lesion is a HCA or a FNH remains unresolved will we administer gadoxetate disodium. Gadobenate dimeglumine is less reliable and therefore should not be used.

In part 2, **chapter 10** we reviewed all current evidence on therapeutic options once a diagnosis of HCA is made. In contrast to other guidelines, we propose use of less stringent follow-up when a L-FABP-negative HCA is diagnosed [5]. In this situation we consider the chance of malignant degeneration to be very low, with the only available evidence derived from case reports [13]. Another important point was our proposal to manage treatable patients in an expert referral centre where expertise is available in treatment options, MRI diagnosis, and pathological sampling.

## Clinical implication

Concerning staging of cervical carcinoma: current international accepted diagnostic work-up for cervical carcinoma is based on clinical examination of the cervix. The most important reason for this is that cervical carcinoma mainly occurs in the tertiary world where MRI is not available. The second reason is that most prognostic and therapeutic studies on cervical carcinoma are based on clinical examination as diagnostic tool. We think that our meta-analysis clearly shows that the info of MRI cannot be obviated anymore. Further prognostics should be correlated with the different MRI based tumour stages and future treatment allocation will preferentially have to be based on MRI instead of clinical examination. It is not defensible anymore to withdraw the info of current state-of-the-art MRI information in the pretreatment set-up of patients with primary cervical carcinoma. However, when state-of-the-art MRI is not available or when a radiologist has little or no experience with MRI of the cervix, clinical examination can be an alternative.

Concerning response evaluation of cervical carcinoma, we propose to include more often MRI in daily settings for response evaluation of cervical carcinoma after radiotherapy. It will not replace the clinical examination, since latter is much cheaper and because small superficial residues will only be visible to the eye.

Concerning MRI of endometriosis, we propose not to use MRI as screening tool for endometriosis. However if MRI is available, it can be used as a reliable technique to diagnose endometriosis if MRI shows hyperintense foci on T1-w imaging.

Concerning chapter 5 until 10, we propose to use MRI with primovist as main diagnostic for differentiation of HCA from FNH. There seems not to be any indication to use the hepatobiliary

phase of multihance. In challenging cases where HCA are hyperintense in the hepatobiliary phase of primovist, radiologist should rely on other typical MRI signs like the atoll-sign and diffuse internal fat. In most circumstances, biopsy will not be needed anymore for diagnosis. However, detection of b-catenin HCA is still difficult non-invasively. Therefore, we propose to still biopsy HCAs if the lesion has atypical findings on MRI and if the lesions are large.

## Future perspectives

In order to familiarize more radiologists with the difficulties of interpreting MRI of cervical cancer, we propose to collect imaging databases with pathologic proof from different centres and to use this for training purposes. Subsequently, interactive courses can be organized with the aim of practising. This is already successfully performed for CT colonography in Amsterdam [14].

Regarding pelvic and liver imaging, an important future direction could be the accumulation of major datasets from our own or other institutes in order to further complement image information. An important condition is consistency of image quality in different institutions. Radiomics and automatic computerized diagnosis will be partly a self-learning tool [15;16]. In this way stronger evidence will be produced on whether this or some other diagnostic tool is still applicable for a certain diagnosis.

In liver imaging, we are working on computer-based models which should eventually help us differentiate malignant from benign lesions. Although this will be of minor importance to specialized radiologists in expert centres, other less specialized radiologists would surely find this tool helpful. In less experienced centres, the most important aim of a local radiologist is to diagnose a lesion as harmless (such as haemangioma, FNH, HCA, etc.) or as suspect and in need of referral to a tertiary centre. A highly-sensitive computerized test would be particularly helpful for these radiologists, while specificity could be more moderate without compromising patient care since false positive results will be identified as such during further workup.

Concerning pelvic imaging, it is not sure if higher Tesla MRIs will further help the radiologist to detect small tumour residue after (chemo) radiation for cervical carcinoma. Also the detection of very small endometriosis implants in the pelvis will remain a challenge. Radiomics may be able to identify which tumours will have better/worse prognosis which may become a focus of research in this field in the near future. Currently, we are evaluating 3D histogram analysis of the DWI data in identifying cervical tumours with significantly worse prognosis. In the future, a more personalized approach will probably indicate which therapy could be optimal.

## References

1. Elliott KS, Borowsky ME, Malka ES, Scudder SA, Leiserowitz GS, Russell AH (2006) Long-term results of maximally aggressive trimodality therapy in a high-risk subset of early-stage cervical cancer patients. *J Reprod Med* 51:383-388
2. Lee JY, Youm J, Kim JW et al (2016) An Alternative Triage Strategy Based on Preoperative MRI for Avoiding Trimodality Therapy in Stage IB Cervical Cancer. *Cancer Res Treat* 48:259-265
3. Hricak H, Gatsonis C, Chi DS et al (2005) Role of imaging in pretreatment evaluation of early invasive cervical cancer: results of the intergroup study American College of Radiology Imaging Network 6651-Gynecologic Oncology Group 183. *J Clin Oncol* 23:9329-9337
4. Khanna M, Ramanathan S, Fasih N, Schieda N, Virmani V, McInnes MD (2015) Current updates on the molecular genetics and magnetic resonance imaging of focal nodular hyperplasia and hepatocellular adenoma. *Insights Imaging* 6:347-362
5. European Association for the Study of the Liver . Electronic address eee (2016) EASL Clinical Practice Guidelines on the management of benign liver tumours. *J Hepatol* 65:386-398
6. Ronot M, Bahrami S, Calderaro J et al (2011) Hepatocellular adenomas: accuracy of magnetic resonance imaging and liver biopsy in subtype classification. *Hepatology* 53:1182-1191
7. Laumonier H, Bioulac-Sage P, Laurent C, Zucman-Rossi J, Balabaud C, Trillaud H (2008) Hepatocellular adenomas: magnetic resonance imaging features as a function of molecular pathological classification. *Hepatology* 48:808-818
8. Yoneda N, Matsui O, Kitao A et al (2012) Beta-catenin-activated hepatocellular adenoma showing hyperintensity on hepatobiliary-phase gadoxetic-enhanced magnetic resonance imaging and overexpression of OATP8. *Jpn J Radiol* 30:777-782
9. Grazioli L, Morana G, Kirchin MA, Schneider G (2005) Accurate differentiation of focal nodular hyperplasia from hepatic adenoma at gadobenate dimeglumine-enhanced MR imaging: prospective study. *Radiology* 236:166-177
10. Paradis V, Champault A, Ronot M et al (2007) Telangiectatic adenoma: an entity associated with increased body mass index and inflammation. *Hepatology* 46:140-146
11. Roux M, Pigneur F, Calderaro J et al (2015) Differentiation of focal nodular hyperplasia from hepatocellular adenoma: Role of the quantitative analysis of gadobenate dimeglumine-enhanced hepatobiliary phase MRI. *J Magn Reson Imaging* 42:1249-1258
12. Neri E, Bali MA, Ba-Ssalamah A et al (2016) ESGAR consensus statement on liver MR imaging and clinical use of liver-specific contrast agents. *Eur Radiol* 26:921-931
13. Stueck AE, Qu Z, Huang MA, Camprecios G, Ferrell LD, Thung SN (2015) Hepatocellular Carcinoma Arising in an HNF-1alpha-Mutated Adenoma in a 23-Year-Old Woman with Maturity-Onset Diabetes of the Young: A Case Report. *Semin Liver Dis* 35:444-449
14. Liedenbaum MH, Bipat S, Bossuyt PM et al (2011) Evaluation of a standardized CT colonography training program for novice readers. *Radiology* 258:477-487
15. Nie K, Shi L, Chen Q et al (2016) Rectal Cancer: Assessment of Neoadjuvant Chemoradiation Outcome based on Radiomics of Multiparametric MRI. *Clin Cancer Res* 22:5256-5264
16. Schlett CL, Hendel T, Weckbach S et al (2016) Population-Based Imaging and Radiomics: Rationale and Perspective of the German National Cohort MRI Study. *Rofo* 188:652-661

# Samenvatting

Dit proefschrift probeert een antwoord te geven op meerdere aspecten van de abdominale beeldvorming door middel van MRI onderzoek bij ziektebeelden welke voornamelijk tot uitsluitend bij vrouwen voorkomen.

Baarmoederhalskanker is een vaak voorkomende ziekte, die oudere maar ook in toenemende mate jongere vrouwen kan treffen. Reeds lang worden verschillende hulpmiddelen gebruikt om in te schatten hoever de baarmoederhalskanker locoregionaal doorgroeit. Dat is belangrijk want bij het ontbreken van lokale doorgroei kan een operatie volstaan, in andere gevallen zal overgegaan worden tot radiotherapie al dan niet met chemotherapie. Doorgaans wordt er een lichamelijk onderzoek uitgevoerd met touche en in speculo visualisatie om dit onderscheid te maken. De waarde van deze techniek is onduidelijk omdat met name de resultaten in de literatuur heel uiteenlopend zijn. Dit hebben we kunnen aantonen met onze meta-analyse, beschreven in **hoofdstuk 2**. Bovendien speelt het ook een rol dat de voorspellende waarde van laterale doorgroei nauwelijks te beoordelen is omdat er meestal geen referentiestandaard is. Het is namelijk zo dat in het geval dat als de gynaecoloog stelt dat er wel laterale doorgroei is, geen referentie standaard is welke kan bevestigen of de gynaecoloog gelijk heeft. In dat geval wordt er geen operatie uitgevoerd en ontbreekt derhalve het pathologisch bewijs en volgt standaard radiotherapie wat kan leiden tot een selectieve verificatie bias. Deze problemen ontstaan ook is in mindere mate tevens aanwezig als we de waarde van MRI willen nagaan ter voorspelling van laterale doorgroei. Maar niettemin kunnen we stellen dat als we alle data over lichamelijk onderzoek en MRI naast elkaar zetten, MRI significant beter scoort. Dit wordt duidelijk als we de verzamelde sensitiviteiten waarbij de laterale uitbreiding wordt uitgesloten en hun betrouwbaarheidsinterval wordt bekeken. Hier blijkt er namelijk geen overlap te zijn. De specificiteiten (hiermee bedoel ik de voorspelling dat er wel laterale doorgroei is) zijn is vergelijkbaar hoog voor beide onderzoekstechnieken. Zoals hierboven uitgelegd is dit laatste echter niet echt altijd hard te maken.

Indien de ziekte lokaal te uitgebreid is lokaal zal overgegaan worden tot radiotherapie-radiotherapie in plaats van een operatie, al dan niet in combinatie met chemotherapie. Tegenwoordig wordt de radiotherapie aangevuld met lokale interne bestraling. Dit geeft doorgaans zeer goede resultaten, doch een klein aantal zal niet reageren op deze standaard therapie. De lokale respons beoordeling wordt klinisch gedaan op het eerste poli bezoek na behandeling. Dit is niet altijd even gemakkelijk door het lokale oedeem en ontsteking van de vagina en de baarmoederhals. Een alternatief is om MRI te gebruiken. Dit hebben wij prospectief onderzocht in een multicenter setting, beschreven in **hoofdstuk 3**. Hieruit komt naar voren dat wanneer MRI gebruikt wordt voor respons beoordeling er altijd een speciale sequentie moet worden toevoegen om het onderscheid tussen oedeem en tumor te kunnen maken, anders is het onvoldoende betrouwbaar met veel vals positieven. Op basis van logistische regressie

analyse hebben we ook kunnen aantonen dat dit MRI protocol gelijkwaardig is aan lichamelijk onderzoek tijdens het poli bezoek. Een voorstel zou kunnen zijn om het lichamelijk onderzoek te vervangen door MRI, indien bijvoorbeeld het lichamelijk onderzoek moeilijk uitvoerbaar is, maar toekomstige studies moeten onze resultaten bevestigen.

In **hoofdstuk 4** hebben we onderzocht of het te voorspellen is of iemand endometriosis heeft op basis van MRI. Endometriosis kan een erg invaliderende ziekte zijn bij jonge vrouwen in de vruchtbare leeftijd, met vaak veel pijnklachten tijdens de menstruatie. Bovendien kan het deels verantwoordelijk zijn voor onvruchtbaarheid. Het is niet altijd even gemakkelijk voor de gynaecoloog om de diagnose te stellen op basis van het verhaal van de patiënt en lichamelijk onderzoek. Soms kunnen deze patiënten ook weinig klachten hebben maar wordt er bij een kijkoperatie een uitgebreide verklevende ziekte gezien. Om bij een patiënt met verminderde vruchtbaarheid zonder oorzakelijke factoren bij de man of de vrouw zeker te zijn dat er geen onderliggende endometriosis aanwezig is, zou je eigenlijk een kijkoperatie moeten uitvoeren. Dit was tot voor kort de standaard procedure. Maar dit is relatief invasief en bovendien duur. Met ons onderzoek hebben we als eerste aangetoond dat met hoge veld MRI aan de hand van een aangepast protocol 3 van de 4 stadia van endometriosis voorspeld kunnen worden. Bovendien blijkt uit deze studie dat als we iets zien op MRI, men redelijk betrouwbaar kan stellen dat de patiënt endometriosis heeft. De studie heeft wel wat pitfalls. Het is een kleine patiëntengroep met een voorafgaandelijke klinische verdenking op endometriosis. Het zou interessant zijn om na te gaan of met deze test endometriosis kan ontdekt worden bij patiënten zonder typische klachten. Vermoedelijk zijn dit patiënten met vooral stadium I endometriosis, en daarbij hebben we juist aangetoond dat MRI hierin vaak te kort schiet. Derhalve is MRI vermoedelijk niet inzetbaar voor screening naar endometriosis.

In deel II eerste **hoofdstuk 5**, hebben we de MRI-kenmerken van verschillende soorten adenomen in de lever geclassificeerd volgens de nieuwe universele indeling als voorgesteld door de groep van Bordeaux.

Adenomen zijn goedaardige tumoren in de lever bij vrouwen voornamelijk in de vruchtbare leeftijd. Op zich benigne, echter kunnen deze tumoren kunnen plots gaan bloeden waarbij meestal een spoedeisende interventie noodzakelijk kan worden. Bovendien kunnen ze in zeldzame gevallen ontwikkelen tot leverkanker.

De groep van Bordeaux heeft enkele jaren geleden een nieuwe classificatie voorgesteld. Het blijkt nu dat 2 subtypes een verhoogde kans hebben om te ontaarden in kanker. Wij hebben aangetoond in dit proefschrift dat er enkele specifieke MRI-kenmerken zijn om deze laesies te identificeren. Verder onderzoek zal moeten uitwijzen of deze kenmerken bruikbaar zijn in een grotere groep van patiënten, bij een zogenaamde externe validatie. We hebben in deze studie ook als eerste het *AtoII*-teken beschreven bij inflammatoire adenomen (een subtype met een verhoogde kans op maligne ontaarding). Deze laesies hebben een dikke dense rand



wat vermoedelijk overeenkomt met peliosis, maar centraal zijn ze even dens als de rest van de lever. Dit geeft het typisch beeld dat ook gezien kan worden bij *atoll* eilanden: een handig trucje dus voor de radioloog en er is al veel naar verwezen in de huidige literatuur.

Een problematische bevinding in de MRI-beeldvorming van adenomen, is dat er een andere levertumor namelijk het FNH (focale nodulaire hyperplasie) er bijna volledig hetzelfde uit kan zien op MRI. Maar deze FNHs zijn volledig ongevaarlijk in tegenstelling tot adenomen. Recentelijk zijn er twee nieuwe MRI-contrastmiddelen op de markt gekomen waarvan ze beweren dat ze betrouwbaar het onderscheid kunnen maken (Multihance en Primovist). Het blijkt dat enige tijdje na de contrasttoediening in een armvat, de lever denser wordt en FNHs ook. Maar adenomen blijkbaar niet. Zó wordt het tenminste gesteld door de fabrikanten. Dit was voornamelijk gebaseerd op een studie van Grazioli et al. in Radiology. In een retrospectieve studie die wij verricht hebben in het Erasmus MC, vonden wij dat dat niet altijd klopt. Met name inflammatoire adenomen kunnen ook denser worden na toediening van Multihance, dit is beschreven in **hoofdstuk 6**. Dit was een volledig nieuwe bevinding, maar andere onderzoeksgroepen bevestigden deze bevindingen. Een andere onderzoeksgroep heeft als alternatief voorgesteld om een kwantificatie methode toe te passen waarbij rekening gehouden wordt met de bevindingen voor contrasttoediening. Wij hebben hierna deze stelling met onze data proberen te valideren, maar wij konden hun methodiek niet reproduceren en dit is. beschreven in **hoofdstuk 8**. In een volgende studie zijn we na gaan kijken hoe het nu zit met Primovist, of daar een gelijkaardig probleem is met inflammatoire adenomen. Er is maar een manier om dat terdege te onderzoeken. Dat is door de patiënten zowel Multihance als Primovist toe te dienen, wel op een ander tijdstip natuurlijk. Hieruit kwam naar voren dat het voornamelijk een Multihance probleem is. Primovist is beduidend beter in deze contrastfase, dit is beschreven in **hoofdstuk 9**.

In de officiële Europese richtlijnen wordt er geen voorkeur voor eender contrastmiddel gegeven, maar met deze data bij de hand, zijn er volgens ons genoeg argumenten om dit aan te passen. Een nadeel van onze studie is wel dat er niet altijd een gouden standaard aanwezig was, een zogenaamd bewijs als het ware welk contrastmiddel gelijk heeft. De uiteindelijke diagnose was in onze studie niet altijd tumor weefsel, maar soms ook een combinatie van verschillende klinische en MRI kenmerken.

Dit is de eerste studie dat aantoont dat de keuze welk contrastmiddel te gebruiken niet vrijblijvend is, en om die reden zijn wij op onze afdeling ook overgegaan van Multihance naar Primovist bij vraagstelling adenoom of FNH.

Aan de hand van de recentste literatuur hebben we nagekeken wat nou het beste beleid is, wanneer bij een jonge vrouw de diagnose gesteld is van een adenoom. In ieder geval wordt dan geadviseerd om te stoppen met hormonen (anticonceptie pil) en als ze dan niet afnemen in omvang dienen ze het beste geëvalueerd te worden in een tertiair referentie centrum voor

eventuele resectie. In **hoofdstuk 10** stellen we voor hoe op basis van de recente literatuur de behandeling best kan plaatsvinden eens de diagnose is gesteld.

Hopelijk hebben we met deze studies nieuwe handvaten kunnen geven bij de beoordeling en beleid van zowel gynaecologische als leveraandoeningen die typisch bij vrouwen voorkomen. In de toekomst zullen deze bevindingen verder gevalideerd moeten worden, en we hopen dan daar ons deel te kunnen bijdragen.

# Chapter 12

List of publications

PhD portfolio

About the author

Dankwoord



# List of publications

## Journal publications

1. **Thomeer MG**, Vandecaveye V, Braun L, Mayer F, Franckena-Schouten M, de Boer P, Stoker J, Van Limbergen E, Buist M, Vergote I, Hunink M, van Doorn H.  
Evaluation of T2-W MR imaging and diffusion-weighted imaging for the early post-treatment local response assessment of patients treated conservatively for cervical cancer: a multicentre study.  
Eur Radiol. 2018 Jun 25.
2. Taimr P, Klompenhouwer AJ, **Thomeer MG**, Hansen BE, IJzermans JNM, de Man RA, de Kneegt RJ.  
Can point shear wave elastography differentiate focal nodular hyperplasia from hepatocellular adenoma.  
J Clin Ultrasound. 2018 Jul;46(6):380-385
3. Toes-Zoutendijk E, Kooyker AI, Elferink MA, Spaander MCW, Dekker E, Koning HJ, Lemmens VE, van Leerdam ME, Lansdorp-Vogelaar I; LECO working group.  
Stage distribution of screen-detected colorectal cancers in the Netherlands.  
Gut. 2018 Sep;67(9):1745-1746
4. Klompenhouwer AJ, Bröker MEE, **Thomeer MG**, Gaspersz MP, de Man RA, IJzermans JNM.  
Retrospective study on timing of resection of hepatocellular adenoma.  
Br J Surg. 2017 Nov;104(12):1695-1703.
5. Klompenhouwer AJ, de Man RA, **Thomeer MG**, IJzermans JN.  
Management and outcome of hepatocellular adenoma with massive bleeding at presentation.  
World J Gastroenterol. 2017 Jul 7;23(25):4579-4586.
6. Bröker MEE, Gaspersz MP, Klompenhouwer AJ, Hansen BE, Terkivatan T, Taimr P, Dwarkasing R, **Thomeer MGJ**, de Man RA, IJzermans JNM.  
Inflammatory and multiple hepatocellular adenoma are associated with a higher BMI.  
Eur J Gastroenterol Hepatol. 2017 Oct;29(10):1183-1188. J Magn Reson Imaging. 2017 Jun 6.
7. **Thomeer MG**, Gest B, van Beek H, De Vries M, Dwarkasing R, Klompenhouwer AJ, De Man RA, IJzermans JN, Braun L.  
Quantitative analysis Of hepatocellular adenoma and focal nodular hyperplasia in the hepatobiliary phase: External validation of LLCER method using gadobenate dimeglumine as contrast agent.  
J Magn Reson Imaging. 2017 Jun 6

8. Toes-Zoutendijk E, van Leerdam ME, Dekker E, van Hees F, Penning C, Nagtegaal I, van der Meulen MP, van Vuuren AJ, Kuipers EJ, Bonfrer JM, Biermann K, **Thomeer MG**, van Veldhuizen H, Kroep S, van Ballegooijen M, Meijer GA, de Koning HJ, Spaander MC, Lansdorp-Vogelaar I; Dutch National Colorectal Cancer Screening Working Group. Real-Time Monitoring of Results During First Year of Dutch Colorectal Cancer Screening Program and Optimization by Altering Fecal Immunochemical Test Cut-Off Levels. *Gastroenterology*. 2017 Mar;152(4):767-775.e2
9. **Thomeer MG**, Broker M, Verheij J, Doukas M, Terkivatan T, Bijdevaate D, De Man RA, Moelker A, IJzermans JN. Hepatocellular adenoma: when and how to treat? Update of current evidence. *Therap Adv Gastroenterol*. 2016 Nov;9(6):898-912
10. Tutein Nolthenius CJ, Boellaard TN, de Haan MC, Nio CY, **Thomeer MG**, Bipat S, Montauban van Swijndregt AD, Essink-Bot ML, Kuipers EJ, Dekker E, Stoker J. Burden of waiting for surveillance CT colonography in patients with screen-detected 6-9 mm polyps. *Eur Radiol*. 2016 Nov;26(11):4000-4010.
11. Klompenhouwer AJ, de Man RA, **Thomeer MG**, Doukas M, IJzermans JN. Diagnostics and treatment of hepatocellular adenomas. *Ned. Tijdschr Geneesk*. 2016(0):D556
12. IJSpeert JE, Tutein Nolthenius CJ, Kuipers EJ, van Leerdam ME, Nio CY, **Thomeer MG**, Biermann K, van de Vijver MJ, Dekker E, Stoker J. CT-Colonography vs. Colonoscopy for Detection of High-Risk Sessile Serrated Polyps. *Am J Gastroenterol*. 2016 Apr;111(4):516-22.
13. Tutein Nolthenius CJ, Boellaard TN, de Haan MC, Nio CY, **Thomeer MG**, Bipat S, Montauban van Swijndregt AD, van de Vijver MJ, Biermann K, Kuipers EJ, Dekker E, Stoker J. Computer tomography colonography participation and yield in patients under surveillance for 6-9 mm polyps in a population-based screening trial. *Eur Radiol*. 2016 Aug;26(8):2762-70.
14. Tutein Nolthenius CJ, Boellaard TN, de Haan MC, Nio CY, **Thomeer MG**, Bipat S, Montauban van Swijndregt AD, van de Vijver MJ, Biermann K, Kuipers EJ, Dekker E, Stoker J. Evolution of Screen-Detected Small (6-9 mm) Polyps After a 3-Year Surveillance Interval: Assessment of Growth With CT Colonography Compared With Histopathology. *Am J Gastroenterol*. 2015 Dec;110(12):1682-90.
15. **Thomeer MG**, Devos A, Lequin M, De Graaf N, Meeussen CJ, Meradji M, De Blaauw I, Sloots CE. High resolution MRI for preoperative work-up of neonates with an anorectal malformation: a direct comparison with distal pressure colostography/fistulography. *Eur Radiol*. 2015 Dec;25(12):3472-9.

16. **Thomeer MG**, E Bröker ME, de Lussanet Q, Biermann K, Dwarkasing RS, de Man R, IJzermans JN, de Vries M.  
Genotype-phenotype correlations in hepatocellular adenoma: an update of MRI findings.  
*Diagn Interv Radiol*. 2014 May-Jun;20(3):193-9.
17. **Thomeer MG**, Steensma AB, van Santbrink EJ, Willemssen FE, Wielopolski PA, Hunink MG, Spronk S, Laven JS, Krestin GP.  
Can magnetic resonance imaging at 3.0-Tesla reliably detect patients with endometriosis? Initial results.  
*J Obstet Gynaecol Res*. 2014 Apr;40(4):1051-8.
18. Dwarkasing RS, Verschuuren SI, Leenders GJ, **Thomeer MG**, Dohle GR, Krestin GP.  
Chronic lower urinary tract symptoms in women: classification of abnormalities and value of dedicated MRI for diagnosis.  
*AJR Am J Roentgenol*. 2014 Jan;202(1):W59-66.
19. **Thomeer MG**, Willemssen FE, Biermann KK, El Addouli H, de Man RA, IJzermans JN, Dwarkasing RS.  
MRI features of inflammatory hepatocellular adenomas on hepatocyte phase imaging with liver-specific contrast agents.  
*J Magn Reson Imaging*. 2014 May;39(5):1259-64.
20. **Thomeer MG**, Gerestein C, Spronk S, van Doorn HC, van der Ham E, Hunink MG.  
Clinical examination versus magnetic resonance imaging in the pretreatment staging of cervical carcinoma: systematic review and meta-analysis.  
*Eur Radiol*. 2013 Jul;23(7):2005-18.
21. van Dam L, de Wijkerslooth TR, de Haan MC, Stoop EM, Bossuyt PM, Fockens P, **Thomeer M**, Kuipers EJ, van Leerdam ME, van Ballegooijen M, Stoker J, Dekker E, Steyerberg EW.  
Time requirements and health effects of participation in colorectal cancer screening with colonoscopy or computed tomography colonography in a randomized controlled trial.  
*Endoscopy*. 2013;45(3):182-8.
22. de Haan MC, de Wijkerslooth TR, Stoop E, Bossuyt P, Fockens P, **Thomeer M**, Kuipers EJ, Essink-Bot ML, van Leerdam ME, Dekker E, Stoker J.  
Informed decision-making in colorectal cancer screening using colonoscopy or CT-colonography.  
*Patient Educ Couns*. 2013 Jun;91(3):318-25.
23. de Wijkerslooth TR, de Haan MC, Stoop EM, Bossuyt PM, **Thomeer M**, van Leerdam ME, Essink-Bot ML, Fockens P, Kuipers EJ, Stoker J, Dekker E.  
Reasons for participation and nonparticipation in colorectal cancer screening: a randomized trial of colonoscopy and CT colonography.  
*Am J Gastroenterol*. 2012 Dec;107(12):1777-83.
24. de Haan MC, **Thomeer M**, Stoker J, Dekker E, Kuipers EJ, van Ballegooijen M.  
Unit costs in population-based colorectal cancer screening using CT colonography performed in university hospitals in The Netherlands.  
*Eur Radiol*. 2013 Apr;23(4):897-907.

25. de Haan MC, Nio CY, **Thomeer M**, de Vries AH, Bossuyt PM, Kuipers EJ, Dekker E, Stoker J.  
Comparing the diagnostic yields of technologists and radiologists in an invitational colorectal cancer screening program performed with CT colonography.  
*Radiology*. 2012 Sep;264(3):771-8.
26. Heijkoop ST, Franckena M, **Thomeer MG**, Boere IA, Van Montfort C, Van Doorn HC.  
Neoadjuvant chemotherapy followed by radiotherapy and concurrent hyperthermia in patients with advanced-stage cervical cancer: a retrospective study.  
*Int J Hyperthermia*. 2012;28(6):554-61.
27. de Wijkerslooth TR, de Haan MC, Stoop EM, Bossuyt PM, **Thomeer M**, Essink-Bot ML, van Leerdam ME, Fockens P, Kuipers EJ, Stoker J, Dekker E.  
Burden of colonoscopy compared to non-cathartic CT-colonography in a colorectal cancer screening programme: randomised controlled trial.  
*Gut*. 2012 Nov;61(11):1552-9.
28. Stoop EM, de Haan MC, de Wijkerslooth TR, Bossuyt PM, van Ballegooijen M, Nio CY, van de Vijver MJ, Biermann K, **Thomeer M**, van Leerdam ME, Fockens P, Stoker J, Kuipers EJ, Dekker E.  
Participation and yield of colonoscopy versus non-cathartic CT colonography in population-based screening for colorectal cancer: a randomised controlled trial.  
*Lancet Oncol*. 2012 Jan;13(1):55-64.
29. Eskens FA, de Jonge MJ, Bhargava P, Isoe T, Cotreau MM, Esteves B, Hayashi K, Burger H, **Thomeer M**, van Doorn L, Verweij J.  
Biologic and clinical activity of tivozanib (AV-951, KRN-951), a selective inhibitor of VEGF receptor-1, -2, and -3 tyrosine kinases, in a 4-week-on, 2-week-off schedule in patients with advanced solid tumors.  
*Clin Cancer Res*. 2011 Nov 15;17(22):7156-63.
30. van Aalten SM, **Thomeer MG**, Terkivatan T, Dwarkasing RS, Verheij J, de Man RA, IJzermans JN.  
Hepatocellular adenomas: correlation of MR imaging findings with pathologic subtype classification.  
*Radiology*. 2011 Oct;261(1):172-81.
31. Farid WR, de Jonge J, Slieker JC, Zondervan PE, **Thomeer MG**, Metselaar HJ, de Bruin RW, Kazemier G.  
The importance of portal venous blood flow in ischemic-type biliary lesions after liver transplantation.  
*Am J Transplant*. 2011 Apr;11(4):857-62.
32. Liedenbaum MH, Bipat S, Bossuyt PM, Dwarkasing RS, de Haan MC, Jansen RJ, Kauffman D, van der Leij C, de Lijster MS, Lute CC, van der Paardt MP, **Thomeer MG**, Zijlstra IA, Stoker J.  
Evaluation of a standardized CT colonography training program for novice readers.  
*Radiology*. 2011 Feb;258(2):477-87.



33. de Wijkerslooth TR, de Haan MC, Stoop EM, Deutekom M, Fockens P, Bossuyt PM, **Thomeer M**, van Ballegooijen M, Essink-Bot ML, van Leerdam ME, Kuipers EJ, Dekker E, Stoker J.  
Study protocol: population screening for colorectal cancer by colonoscopy or CT colonography: a randomized controlled trial.  
BMC Gastroenterol. 2010 May 19;10:47.
34. Liedenbaum MH, van Rijn AF, de Vries AH, Dekker HM, **Thomeer M**, van Marrewijk CJ, Hol L, Dijkgraaf MG, Fockens P, Bossuyt PM, Dekker E, Stoker J.  
Using CT colonography as a triage technique after a positive faecal occult blood test in colorectal cancer screening.  
Gut. 2009 Sep;58(9):1242-9.
35. **Thomeer MG**, Pattynama PM, Hartmann IJ, Kieft GJ, Van Strijen MJ.  
High incidence of isolated subsegmental pulmonary emboli on multi-slice spiral CT: a comparative clinical study.  
Thromb Haemost. 2006 May;95(5):914-5.
36. Van Buuren H, Brink MA, Van Der Werf S, Meijssen M, **Thomeer MG**, Fritscher-Ravens A.  
Imaging Characteristics of Autoimmune Pancreato-Cholangitis - Results of a Multicenter Study Group.  
Gastrointestinal Endoscopy. 2006 April Vol 63, issue 5: AB286
37. den Bakker MA, **Thomeer M**, Maat AP, Groeninx van Zoelen CE.  
Life-threatening hemoptysis caused by chronic idiopathic pulmonary hilar fibrosis with unilateral pulmonary vein occlusion.  
Ann Diagn Pathol. 2005 Dec;9(6):319-22.
38. Nieuwenhuis RF, Ossewaarde JM, Götz HM, Dees J, Thio HB, **Thomeer MG**, den Hollander JC, Neumann MH, van der Meijden WJ.  
Resurgence of lymphogranuloma venereum in Western Europe: an outbreak of Chlamydia trachomatis serovar I2 proctitis in The Netherlands among men who have sex with men.  
Clin Infect Dis. 2004 Oct 1;39(7):996-1003.
39. **Thomeer M**, Carbone I, Bosmans H, Kiss G, Bielen D, Vanbeckevoort D, Van Cutsem E, Rutgeerts P, Marchal G.  
Stool tagging applied in thin-slice multidetector computed tomography colonography.  
J Comput Assist Tomogr. 2003 Mar-Apr;27(2):132-9.
40. Bielen D, **Thomeer M**, Vanbeckevoort D, Kiss G, Maes F, Marchal G, Rutgeerts P.  
Dry preparation for virtual CT colonography with fecal tagging using water-soluble contrast medium: initial results.  
Eur Radiol. 2003 Mar;13(3):453-8.
41. **Thomeer M**, Bielen D, Vanbeckevoort D, Dymarkowski S, Gevers A, Rutgeerts P, Hiele M, Van Cutsem E, Marchal G.  
Patient acceptance for CT colonography: what is the real issue?  
Eur Radiol. 2002 Jun;12(6):1410-5.

42. Kiss G, Van Cleynenbreugel J, **Thomeer M**, Suetens P, Marchal G.  
Computer-aided diagnosis in virtual colonography via combination of surface normal and sphere fitting methods.  
Eur Radiol. 2002 Jan;12(1):77-81.
43. **Thomeer M**, Vanbeckevoort D, Bielen D, Beenen L, Gevers A, Rutgeerts R, Marchal G.  
Virtual colonoscopy: a new screening tool for colorectal cancer?  
JBR-BTR. 2001 Aug;84(4):155-63.
44. **Thomeer M**, Vanbeckevoort D, Van Breuseghem I, Petré C, De Vuysere S, Coenegrachts K, Miserez M.  
Appendicitis after appendectomy: CT diagnosis.  
Eur Radiol. 2000;10(4):674-6.

## Chapter

1. Medical Image Computing and Computer-assisted Intervention. Computer aided diagnosis for virtual colonography. Gabriel Kiss, Van Cleynenbreugel Johan, Thomeer Maarten, Paul Suetens, Marchal Guy.  
Multislice CT: a practical Guide: Virtual endoscopy. Didier Bielen, Dirk Vanbeckevoort, Maarten Thomeer , Marc Peeters  
Vision,modeling, and Visualization 2002, Proceedings ,nov 20-22, 2002. P 27-34

# PhD Portfolio

Name PhD student: **Maarten G J Thomeer**  
 Erasmus MC Department: Radiology and Nuclear Medicine  
 PhD period: 2011-2018

## 1 Poster presentations

Vanhooymissen I, Thomeer MG, Gest B, van Koevorden S, Braun L, Dwarkasing R, De Man R, IJzermans J. Intra-patient comparative MRI study of gadobenate dimeglumine and gadoxetate disodium in the differentiation of hepatocellular adenoma and focal nodular hyperplasia. Radiologendagen 11 to 12 may 2017 **ECTS 3**

Miclea R, Gest B, Thomeer MG, Willemssen F, Dwarkasing R. Radiological – pathological correlation of hepatocellular adenoma (HCA) and focal nodular hyperplasia (FNH) with atypical clinical or radiological features on MRI with liver-specific contrast agent. Radiologendagen 11 to 12 may 2017 **ECTS 1**

Van Doorn H, Thomeer MG. Clinical examination versus magnetic resonance imaging in the pretreatment staging of cervical carcinoma: systematic review and meta-analysis. The 2<sup>nd</sup> IGCS Regional Meeting on Gynecologic Cancers, Bali. 11 to 13 April 2013 **ECTS 1**

## 2 Oral presentations

Thomeer MG, Vanhooymissen I, Gest B, van Koevorden S, Braun L, Dwarkasing R, De Man R, IJzermans J. Intra-patient comparative MRI study of gadobenate dimeglumine and gadoxetate disodium in the differentiation of hepatocellular adenoma and focal nodular hyperplasia. **ECTS 4**

Thomeer MG, Gest B, van Beek H, De Vries M, Dwarkasing R, Braun L. Quantitative analysis of hepatocellular adenoma and focal nodular hyperplasia in the hepatobiliary phase: validation of SIR and LLCER method using gadobenate dimeglumine contrast agent. 8<sup>th</sup> International Liver Forum 2017, 19-21 oktober 2017 **ECTS 3**

Thomeer MG: radiologie bij benigne leverhaarden. Congres Nederlandse Vereniging van Gastro-enterologie. Veldhoven, 6 en 7 oktober 2016 **ECTS 3**

Thomeer MG en Merel Scheurkogel: Update on Hepatic Adenoma and FNH, including proper use of Hepatobiliary MR Contrast for diagnosis. Sandwich cursus abdominale radiologie, Ede, 8 en 9 juni 2016 **ECTS 4**

Vanhooymissen I, Thomeer MG, Gest B, van Koevorden S, Braun L, Dwarkasing R, De Man R, IJzermans J. Intra-patient comparative MRI study of gadobenate dimeglumine and gadoxetate disodium in the differentiation of hepatocellular adenoma and focal nodular hyperplasia. ECR 2016. **ECTS 3**

Thomeer MG High resolution MRI for preoperative work-up of neonates with an anorectal malformation: a direct comparison with rectofistulography using surgical findings as reference standard. European Congress of Radiology, Vienna, 4 to 8 maart 2015. **ECTS 3**

Thomeer MG MRI features of inflammatory hepatocellular adenomas on hepatocyte phase imaging with liver-specific contrast agents. European Congress of Radiology, Vienna, 4 to 8 maart 2015. **ECTS 3**

Thomeer MG: MRI beeldvorming bij endometriosis. Aios gynaecologie onderwijsdag, Rotterdam, 6 februari 2015. **ECTS 3**

Thomeer MG, Wat wil de gynaecoloog zeker in het CT-verslag hebben staan. Bijeenkomst samenscholing gynecologen en radiologen: een kijkje in elkaars keuken en leren van elkaar, Rotterdam, 18 mei 2015. **ECTS 2**

Thomeer MG High resolution MRI for preoperative work-up of neonates with an anorectal malformation: a direct comparison with rectofistulography using surgical findings as reference standard. Radiologendagen 11 to 12 september 2014 **ECTS 2**

Thomeer MG Hepatocellular adenoma: to treat or not to treat? A comprehensive analysis of the current literature. Radiologendagen 11 to 12 september 2014 **ECTS 2**

De Graaf N, Thomeer MG: High resolution MRI for preoperative work-up of neonates with an anorectal malformation: a direct comparison with rectofistulography using surgical findings as reference standard. European Society for Paediatric Research, Amsterdam, 2 to 6 Juni 2014 **ECTS 1**

Sloots C, Thomeer MG: High resolution MRI for preoperative work-up of neonates with an anorectal malformation: a prospective study comparing with fluoroscopic studies and surgical findings as reference standard. European Paediatric Surgeons' Association, Dublin, 18 to 21 juni 2014 **ECTS 1**

Thomeer MG: Beeldvorming bij ovariumcarcinoom met casuïstiek. Aios gynaecologie onderwijsdag, Rotterdam, 16 mei 2014 **ECTS 2**

Thomeer MG Case presentation anorectal malformation. Annual Pediatric Colorectal Workshop, Rotterdam, 9 to 12 october 2013 **ECTS 1**

Thomeer MG: Afscheidssymposium Maria vd Burg ECTS 1

Thomeer MG: Radiologie voor dummies. Aios gynaecologie onderwijsdag, Rotterdam, 15 maart 2013 **ECTS 1**

Van Buuren H, Brink MA, Van Der Werf S, Meijssen M, Thomeer M, Fritscher-Ravens A. Imaging Characteristics of Autoimmune Pancreato-Cholangitis - Results of a Multicenter Study Group. MG,Fritscher-Ravens A. Gastrointestinal Endoscopy. 2006 April Vol 63, issue 5: AB286 **ECTS 1**

Thomeer MG: First liver MRI workshop Rotterdam, Rotterdam, 4-5 oct 2018 **ECTS 4**

Thomeer MG: Aios gynaecologie onderwijsdag, 15 maart 2013 **ECTS 2**

## Cursussen

NIHES (biostatistics for clinicians), Rotterdam, 2011 **ECTS 4**

Biostatistical methods (NIHES), Rotterdam, 2013 **ECTS 5**

Teach the teacher III course, Rotterdam 2015 **ECTS 3**



# Curriculum vitae

## Maarten Guillaume Josephus Thomeer

Maarten Thomeer is geboren op 10 november 1970 te Seria in Brunei.

Hij heeft 3 broers, en reist in zijn jeugd met het gezin naar Angola. Ze verblijven daar tot in 1978, en verhuizen dan naar Hoeilaart bij Brussel. In 1989 haalt hij zijn diploma middelbaar onderwijs richting Latijn-Wiskunde aan het Sint-Jozefs-College te Brussel. Hij start hetzelfde jaar de opleiding Geneeskunde aan de Katholieke Universiteit van Leuven om af te studeren in 1996 als arts. Tijdens de opleiding volgt hij meerdere stages, waaronder een te Dakar in Senegal. Hij studeert af als 'geneesheer-specialist in de röntgendiagnose' aan de Universitaire Ziekenhuizen te Leuven in 2002 bij toenmalig diensthoofd Professor Guy Marchal. Vanaf dat jaar werkt hij in dienst van de afdeling Radiologie van het Erasmus Universitair Medisch Centrum te Rotterdam in Nederland. Hij doet daar nu voornamelijk abdominale en thorax radiologie en leidt mede assistenten op. Hij is getrouwd met Christine Groeninckx. Samen hebben ze 4 kinderen: Maartje, Nikolaas, Lineke en Noortje. Het gezin woont nu in Essen, België.





# Dankwoord

*Ja jongens; Het is eindelijk af!*

Jullie zijn met velen betrokken bij dit boekje en de verdediging ervan. Ik ben me ervan bewust dat ik in dit dankwoord mensen zal vergeten te noemen, daarom wil ik mijn dankbaarheid uiten aan allen die meegewerkt hebben aan deze thesis, zowel direct als indirect. Los hiervan wil ik bepaalde personen nog meer in de bloemetjes zetten.

Geachte radiologische leden van de commissie, beste Myriam Hunink, Vincent Vandecaveye, Jaap Stoker (tevens kleine commissie) en Regina Beets-Tan (kleine commissie). Ik heb jullie gekozen vanwege jullie ontegensprekelijke uitgebreide klinische en wetenschappelijke bijdrage de laatste jaren op gebied van abdominale MRI zowel nationaal en internationaal. Ik ben verheugd U te mogen opponeren.

Geachte niet-radiologische leden van de commissie, beste Lena van Doorn, Jan IJzermans (tevens kleine commissie), Rob de Man, en Chris Verslype. Dank dat jullie in de commissie willen zetelen. Met jullie uitgebreide klinische achtergrond hoop ik in mijn verdediging ook jullie te kunnen overtuigen van de steeds belangrijkere positie van de diagnostische beeldvorming in de dagdagelijkse medische praktijk.

Vooreerst bedankt professor Hunink. Liefste Myriam, het is een hele eer om u te hebben als promotor. Ik ben u heel dankbaar voor het geloof in onze onderzoekslijn, die we samen verder uitgestippeld hebben. Soms werd het zo complex, dat ik blij was dat u weer kwam met een creatieve oplossing. Hoewel u een drukke agenda heeft, hebt u altijd tijd vrij kunnen maken voor immer verhelderende discussies.

Ik zou ook mijn dankbaarheid willen uiten aan co-promotor Dr. van Doorn. Liefste Lena, wat een lang traject hebben we achter onze rug. Jij en professor Curt Burgers, waren er vanaf het begin bij. Jullie hebben me overtuigd dat een gespecialiseerde radioloog een wezenlijk onderdeel vormt van de Multidisciplinaire Gynaecologie Oncologie besprekingen. We hebben beide een hart voor onderzoek, maar bovenal sta je bekend om je lovenswaardige betrokkenheid bij je patiënten. Ik heb dat ook al meerdere keren mogen ervaren zowel in mijn professionele als in mijn privé leven. Daarvoor ben ik je heel dankbaar.

Beste professor Krestin, beste Gabriel. Het is alweer van 2002 geleden dat u mij vroeg stafleid te worden in de abdominale sectie. De abdomen sectie was toen in fragiel evenwicht maar is nu uitgegroeid tot een sterke 7 koppige groep met elk zijn specifieke kennis en liefde voor

het vak van de abdominale radiologie. Uw organisatiekwaliteiten en geduld hebben onze afdeling gebracht tot waar het nu hoort te staan.

Beste Professoren Jan IJzermans en Rob de Man, bedankt voor jullie support. In het begin heb ik getwijfeld om mijn onderzoeksdata over adenomen te includeren in mijn boekje. Het werd me echter gauw duidelijk dat de hepatobiliaire groep van Rotterdam een wezenlijk begrip is in Nederland en omstreken. Jullie zijn daar in groten delen voor verantwoordelijk. Ik kon het daarom niet laten onze gezamenlijke onderzoeksdata over adenomen te implementeren in mijn boekje.

Dr. Steensma, liefste Anneke. Het zijn altijd al vruchtbare onderlinge discussies geweest zowel tijdens als buiten de Werkgroep Bekkenbodem. Ik waardeer sterk je kennis en praktisch inzicht in complexe bekkenbodem pathologie. Je was altijd aanspreekbaar en makkelijk toegankelijk, daarvoor dank ik je.

Loes Braun, voor jou ook een ereplekje. Vele lange sessies hebben we samen doorgebracht om op een of andere manier waardevolle statistische analyses te kunnen maken, en dat vooral voor de cervixstudie, maar ook voor de papers over adenomen.

Een hééééél belangrijk dankwoord voor iedereen van het trialbureau, en nog in het bijzonder Laurens Groenendijk, Caroline van Bavel en Brigitte Hofstee. Het staat buiten kijf dat de cervix studie tot niets geleid zou hebben als jullie er niet altijd waren. Zowel voor de inclusies als voor het vele papierwerk kon ik op jullie rekenen. Ik was soms veeleisend, maar gelukkig konden jullie het allemaal relativeren.

Ik zou ook graag Nuts-Ohra willen bedanken voor de financiële ondersteuning van de multicenter studie over het cervix carcinoom.

Mijn collega's van de Abdominale sectie. Met name Roy Dwarkasing, Nanda Krak, Ivo Schoots, Barbara Simons, Marianne de Vries, en François Willemssen en ook de fellows Julia Galuzina, Marzena Kulawska, Marion van Gils en Inge Vanhooymissen. De Abdominale radiologie is het epicentrum van onze afdeling alleen al op basis van klinische productiecijfers, met veel onderzoeksmogelijkheden, we mogen de boot daarom niet missen. Maar zoals een van ons ooit zei: 'eerst even de werklijst leegtrekken'. Ik hoop dat we in de toekomst meer tijd kunnen vrijmaken voor research projecten en onderwijs.

Paranymfen: Arlette Odink en Adriaan Moelker, Ik heb jullie mede mogen opleiden en ben nu vereerd dat jullie ook nu aan mijn zijde staan. Jullie zijn altijd zo vriendelijk en toegankelijk geweest, eventjes brainstormen of gewoon bijpraten en dan waren we weer bij.

Oud-opleider Professor Emeritus Marchal, beste Guy dank van harte om mij de weg te wijzen tijdens mijn opleiding en dank om het geloof in mijn kunnen. Van u heb ik de passie voor het vak gekregen.

Alle Coauteurs bedankt voor jullie input aan dit project, met name:

Jaap Stoker, Peter de Boer, Vincent Vandecaveye, Frederik De Keyzer, Martine Franckena-Schouten, Frenchey Mayer, Erik Van Limbergen, Ignace Vergote, Marrije Buist, Joop Laven, Joanne Verheij, Mirelle Broker, Michael Doukas, Turkan Terikavatan, Diederick Bijdevaate, Adriaan Moelker, Katharina Biermann, Quido de Lussanet en Marianne De Vries, Evert van Santbrink, Piotr Wielopolski, Sandra Spronk, Haroun El Addouli, Cees Gerestein, Julia Klompenhouwer, Sebastiaan Van Koevorden, Inge Vanhooymissen, Bibiche Gest, Sanne van Aalten, Hermen van Beek, en Els van der Ham.

Rob Voogt, en Arjan Koppert, ik weet niet waarom ik jullie hier vernoem, maar jullie zullen het waarschijnlijk wel kunnen appreciëren. Dank ook voor jullie vriendschap.

Mijn vader en broers Michiel, Roel en Koentje: sinds het overlijden van moeders vijf jaar geleden heb ik mogen ondervinden dat onze band heel hecht geworden is, we kunnen altijd bij elkaar terecht voor onze zielenroerselen en kleine familie perikelen; ik denk dat moeders ons stiekem in de gaten houdt en beschermd.

Mijn kinders, Maartje, Nikolaas, Lineke en Noortje, jullie zullen op de eerste rij zitten bij mijn promotie, en samen met mama staan jullie ook vooraan in mijn leven.

En als laatste: allerliefste Christine, mijn alles, het is geweldig hoe je altijd in mij gelooft in al die jaren. Het promotiestuk is klaar, maar de drukte zal vermoedelijk niet direct overgaan. Toch lukt het ons om steeds terug te vallen op onze liefde voor elkaar en de kinders, inclusief Jip, Janneke, Mos en Mopje :de 4 katten, Lola, Fluffie 1, Fluffie 2, Brownie, Brownie Jr, Pinkie, Dommie, Pim, Bos, Sneeuw witje, Marshmallow, Driepoot: de cavia's en natuurlijk Pluisje en Roosje: de konijnen. En niet te vergeten de gestorven cavia in het vriesvak die nu echt wel dringend begraven moet worden.





