

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/53684>

Please be advised that this information was generated on 2019-06-19 and may be subject to change.

Reliability of the Roenigk Classification of Liver Damage After Methotrexate Treatment for Psoriasis

A Clinicopathologic Study of 160 Liver Biopsy Specimens

Maartje A. M. Berends, MD; Martijn G. H. van Oijen, MSc; Josje Snoek, MSc; Peter C. M. van de Kerkhof, MD, PhD; Joost P. H. Drenth, MD, PhD; J. Han van Krieken, MD, PhD; Elke M. G. J. de Jong, MD, PhD

Objective: To determine the interobserver reliability of the Roenigk score as a classification system of liver damage and its possible consequences for clinical practice.

Design: Retrospective study.

Setting: Academic research.

Patients: One hundred sixty liver biopsy specimens from patients with psoriasis receiving methotrexate treatment were rereviewed and analyzed blindly by an experienced pathologist with an interest in liver pathologic conditions.

Main Outcome Measure: Interobserver variation was evaluated using κ statistics.

Results: A high concordance was present in the evaluation of the Roenigk grade of fibrosis (weighted $\kappa=0.73$; 95% confidence interval, 0.63-0.83). Agreement was good regarding the number of biopsy specimens for patients whose clinical management should be changed ($\kappa=0.71$; 95% confidence interval, 0.56-0.87).

Conclusion: The Roenigk classification in the assessment of liver fibrosis is a reliable scoring system.

Arch Dermatol. 2007;143(12):1515-1519

HEPATIC FIBROSIS AND CIRRHOSIS represent a consequence of methotrexate treatment in patients with psoriasis.¹⁻⁶ Therefore, the assessment of liver damage is essential in the clinical management of these patients. Sequential liver biopsies followed by Roenigk grading by a pathologist are the mainstay in the assessment of the stage and degree of liver damage.⁷⁻⁹ Unfortunately, liver biopsies may be associated with sampling error, potential complications, and interobserver variability.^{1,7,9-13}

Methotrexate-associated liver damage in patients with psoriasis is graded according to the Roenigk classification.¹ The results of the Roenigk scoring system should be reproducible, with little interobserver error.

The Roenigk classification was developed by the Psoriasis Task Force (led by dermatologists), is based on clinical observations, and has been recommended in the American Academy of Dermatology guidelines for monitoring methotrexate-induced liver injury.^{1,9,14} However, the Roenigk grading system is subjective, including some features (such as nuclear

pleomorphism) of unclear significance, and is insensitive to small changes, particularly when assessing fibrosis.^{1,15,16} Although scoring seems to consider changes such as steatosis and inflammation, their presence or absence has no weight in the allocation to more advanced grades. The system categorizes all biopsy specimens with more than minimal fibrosis as advanced fibrosis (Roenigk grade 3b) and overestimates the degree of histologic change. Accurate assessment is essential because misclassification of pathologic changes affects clinical management. For example, if the degree of fibrosis is upgraded from none (Roenigk grade 2) to mild (Roenigk grade 3a), guidelines call for a second liver biopsy within 6 months instead of a 1.5-g cumulative dose of methotrexate.⁹ In the case of Roenigk grade 3b or 4, the guidelines recommend discontinuation of therapy.

In some European countries, the number of liver biopsies has declined for several reasons. The use of the aminoterminal propeptide of type III procollagen (PIIINP) has reduced the number of liver biopsies in Scandinavia and in Great Britain. Until recently, no noninvasive method

Author Affiliations:

Departments of Dermatology (Drs Berends, van de Kerkhof, and de Jong), Gastroenterology and Hepatology (Mr van Oijen, Ms Snoek, and Dr Drenth), and Pathology (Dr van Krieken), Radboud University Nijmegen Medical Centre, Nijmegen, the Netherlands.

Table 1. Roenigk Classification System

Fatty Change	Nuclear Pleomorphism	Fibrosis	Necroinflammation	Roenigk Grade ^a
Mild or none	Mild or none	None	With or without mild portal inflammation	1
Moderate or severe	Moderate or severe	None	Moderate or severe portal inflammation	2
With or without	With or without	Mild (fibrosis extending into acini)	With or without	3a
With or without	With or without	Moderate or severe	With or without	3b
With or without	With or without	Cirrhosis	With or without	4

^aSee the "Pathologic Examination" subsection of the "Methods" section for an explanation of Roenigk grades.

has been available that could completely replace the liver biopsy. In the case of a persistently elevated PIIINP, liver biopsy is still advised.^{11,13,17} Given the critical nature of this tool, we evaluated interobserver variation using a sample of 160 liver biopsy specimens from methotrexate-treated patients with psoriasis.

METHODS

PATIENTS

We evaluated interobserver variation between several different pathologists with an interest in liver pathologic routine clinical practice and 1 of us (J.H.v.K.) in the assessment of the histopathologic degree of liver damage according to the Roenigk scale in patients with psoriasis receiving methotrexate treatment. All pathologists were trained at the same department of pathology at the same hospital.

One hundred twenty-five patients with psoriasis had undergone 278 liver biopsies while receiving methotrexate treatment from November 1, 1976, to December 31, 2005. We excluded biopsies performed before December 31, 1995, because these specimens were unavailable for review. In addition, 9 biopsy specimens were excluded from analysis (6 because they were unavailable from the department's archives, 1 was too small to evaluate the degree of fibrosis, and 2 because the van Gieson-stained slide was unavailable). One hundred sixty liver biopsy specimens from 95 patients were reexamined independently by 1 of us (J.H.v.K.) who was blinded to the clinical details of the patients. Liver biopsy specimens were graded according to the Roenigk classification (**Table 1**).¹

HISTOLOGIC EXAMINATION

Percutaneous liver biopsy was performed via a right intercostal approach using local lidocaine hydrochloride anesthesia. The biopsy specimen was immersed in 2% formaldehyde and was subsequently fixed with paraffin. Hematoxylin-eosin-stained sections of liver tissue were examined for steatosis, nuclear variability, hepatocyte necrosis, and lobular and portal tract inflammation. A van Gieson stain for collagen was used to assess for the expansion of the portal tracts and for the presence of pericellular and perivenular fibrosis.

PATHOLOGIC EXAMINATION

All liver biopsy specimens, sampled as part of the monitoring process of methotrexate-induced hepatic injury, were graded according to the Roenigk classification. A description of the Roenigk classification is given in Table 1. Roenigk grade 1 indicates normal tissue with no fibrosis, no or mild portal inflammation, and no or mild fatty changes and nuclear pleomorphism. Grade 2 indicates no fibrosis and moderate or severe

fatty changes, nuclear pleomorphism, and portal inflammation. Grade 3a indicates mild fibrosis, portal fibrotic septa, extension into the lobuli, and portal tract enlargement. Grade 3b indicates moderate or severe fibrosis. Grade 4 indicates cirrhosis, regenerating noduli, and bridging of the portal tracts.

STATISTICAL ANALYSIS

To judge the degree of interobserver agreement, we calculated weighted κ statistics for the 5-point Roenigk scale. For analysis of clinical consequences, we dichotomized the Roenigk score into "no changes of treatment necessary" (Roenigk grades 1 and 2) and "change of treatment necessary" (Roenigk grades 3a, 3b, and 4) for all observations. For agreement of the dichotomized data, we again used κ statistics. Interpretation of the κ statistics was performed using the scale described by Landis and Koch,¹⁸ in which κ statistics less than 0.4 indicate poor agreement, κ statistics between 0.4 and 0.6 indicate moderate agreement, between 0.6 and 0.8 good agreement, and greater than 0.8 indicate excellent agreement.¹⁹

To visualize agreement, we plotted a Bland-Altman curve for the 5-point Roenigk score. Using a 2-sided *t* test, we tested whether the overall mean differences differed statistically significantly from 0. All analyses were undertaken using statistical software (SAS version 8.2; SAS Institute Inc, Cary, North Carolina).

RESULTS

DEMOGRAPHICS

Ninety-five patients with psoriasis (44 female and 51 male) underwent a liver biopsy between December 31, 1995, and December 31, 2005. The maximum prescribed weekly dosage of methotrexate was 12.5 mg (range, 7.5-25 mg). Patients received a median cumulative methotrexate dose of 2051 mg (range, 119-20 235 mg) during a median follow-up period of 202 weeks (range, 20-1763 weeks).

LIVER BIOPSY SPECIMENS

The concordance between the Roenigk grades as scored during routine assessment and at subsequent scoring by the second pathologist was high (weighted $\kappa=0.73$; 95% confidence interval, 0.63-0.83). The agreement was higher for biopsy specimens that were graded as Roenigk grade 1, which was the most common Roenigk score (**Figure**). The mean difference was 0.03 and did not significantly differ from 0 ($P>.05$). Among liver biopsy specimens that resulted in a change of clinical management (Roenigk grades 3a, 3b, and 4), we like-

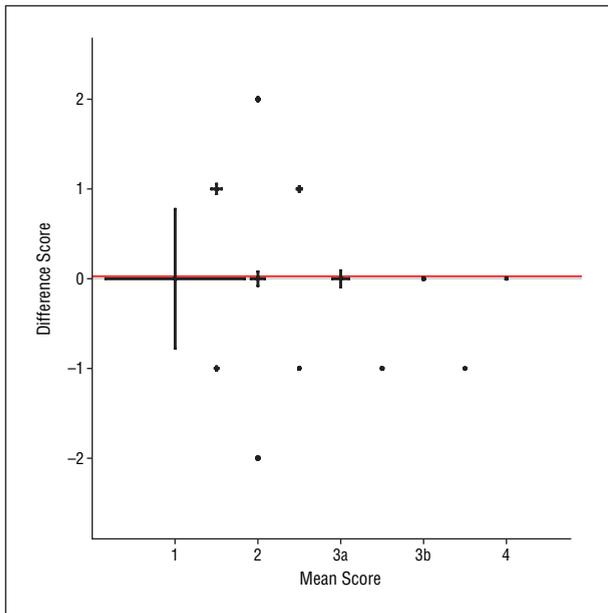


Figure. Bland-Altman curve. The graph shows the absolute difference between the initial and subsequent scores (on the y-axis) against the mean of both scores (on the x-axis) for each observation. In the graph, the size of the crosses indicates how often a pair of observations is found. The red line indicates the 0.03 overall mean difference between the initial and subsequent scores.

wise observed a good correlation ($\kappa=0.71$; 95% confidence interval, 0.56-0.87).

The initial routine examination had graded 113 liver biopsy specimens as Roenigk grade 1, 21 as grade 2, 22 as grade 3a, 3 as grade 3b, and 1 as grade 4 (**Table 2**). After reexamination of all liver biopsy specimens by the second pathologist, 118 were graded as grade 1, 18 as grade 2, 18 as grade 3a, 4 as grade 3b, and 2 as grade 4.

Six liver biopsy specimens originally scored as Roenigk grade 1 were scored differently by the second pathologist (3 as grade 2 and 3 as grade 3a). Ten liver biopsy specimens originally scored as grade 2 were subsequently scored differently (2 were upgraded to grade 3a, while 8 were downgraded to grade 1). Nine liver biopsy specimens originally scored as grade 3a were scored differently by the second pathologist (2 as grade 3b, 4 as grade 2, and 3 as grade 1). Finally, 1 liver biopsy specimen originally scored as grade 3b was subsequently upgraded to grade 4.

CLINICAL CONSEQUENCES OF DIFFERENCES IN SCORING

Fourteen biopsy specimens were downgraded or upgraded to such an extent that it would have affected clinical management (**Table 3**). Seven biopsy specimens graded as Roenigk grade 3a were downgraded by the second pathologist to grade 2 or 1. Because of the original grade, follow-up biopsies in 3 patients were performed after 5 to 10 months, and methotrexate treatment in 1 patient was discontinued after 5 months. Three biopsy specimens were upgraded by the second pathologist from grade 1 to 3a, 2 biopsy specimens from grade 2 to 3a, and 2 biopsy specimens from grade 3a to 3b. In the latter 2 cases, this assessment resulted in follow-up

Table 2. Numbers of Slides Assigned to the Roenigk Grades by the Pathologists^a

Roenigk Grade	First Pathologist	Second Pathologist
1	113	118
2	21	18
3a	22	18
3b	3	4
4	1	2

^aSee the "Pathologic Examination" subsection of the "Methods" section for an explanation of Roenigk grades.

biopsies after 9 and 14 months that demonstrated histologic findings corresponding to grade 3a. One of these patients is still being treated with methotrexate. In the other patient, methotrexate treatment was continued, and 2 more biopsies were performed. Both biopsy specimens demonstrated histologic findings corresponding to grade 3a. Methotrexate treatment was discontinued for an unknown reason.

COMMENT

Our objectives were to determine the interobserver reliability of the Roenigk score as a classification system of methotrexate-induced liver damage and to assess the consequences for clinical practice. The results of this study show that the Roenigk classification is a reliable scoring system for the assessment of liver fibrosis.

The study revealed high concordance between the first and second observations. Also, there was good agreement on biopsy specimens that resulted in a Roenigk grade that necessitated change of clinical management (biopsy specimens with grades 3a, 3b, and 4). Although only a small percentage of the biopsy specimens was scored differently by the second pathologist, it would have resulted in a clear change in the clinical decisions made. Grade 3a requires more frequently performed liver biopsies (within 6 months instead of a 1.5-g cumulative dose of methotrexate), and grades 3b and 4 necessitate interruption and cessation of methotrexate treatment.⁹

Periodic liver biopsies are recommended by international guidelines^{4,9,14} on methotrexate treatment in patients with psoriasis, and the Roenigk score has been recommended in the American Academy of Dermatology guidelines^{4,9,14} to classify methotrexate-induced liver damage. However, the Roenigk scale has not been validated or used (to our knowledge) in the evaluation of any other liver disease.¹ Furthermore, the Roenigk scale is subjective and is insensitive to small changes, particularly when assessing fibrosis.^{1,16} Scoring systems for liver damage such as the Metavir, Scheuer, and Ishak classifications are well established for hepatitis C and for some forms of nonviral hepatitis; these scoring systems are more sensitive to small changes, and studies^{1,20-24} demonstrated good agreement for fibrosis assessment. As far as we know, there are no studies evaluating the validity and interobserver reliability of the Roenigk score. However, 2 studies compare the Roenigk classification with other scoring systems. One study¹⁵ compares the Roenigk score with a

Table 3. Slides Scored Differently by the Pathologists

Roenigk Grade ^a		No. of Slides Scored Differently	Upgrade or Downgrade	Clinical Consequences
First Pathologist	Second Pathologist			
1	2	3	Upgrade	No
1	3a	3	Upgrade	Yes
2	1	8	Downgrade	No
2	3a	2	Upgrade	Yes
3a	1	3	Downgrade	Yes
3a	2	4	Downgrade	Yes
3a	3b	2	Upgrade	Yes
3b	4	1	Upgrade	No

^aSee the "Pathologic Examination" subsection of the "Methods" section for an explanation of Roenigk grades.

semiquantitative histologic scoring system for the evaluation of hepatic fibrosis in patients with rheumatoid arthritis treated with methotrexate. A statistically significant correlation was found between the 2 classification systems, but the semiquantitative histologic scoring system was much more sensitive than the Roenigk score for the assessment of hepatic fibrosis. Another study¹ compared 3 scoring systems for the evaluation of hepatic fibrosis in patients with psoriasis treated with methotrexate. The Roenigk classification was compared with the Scheuer and Ishak scoring systems and seemed to correlate poorly with both systems.

The already described simplification of the Roenigk classification may have improved the interobserver reliability in our study. This raises the question of whether the Roenigk classification is the best-designed scoring system to classify methotrexate-induced liver injury. However, that was not the objective of our study. The Roenigk classification is used by many pathologists to classify methotrexate-induced liver fibrosis. In this study, it is shown that the interobserver reliability is good.

In 8% of the liver biopsy specimens, a different clinical decision would have been made based on disagreement between the first and second observers. When this leads to more frequently performed liver biopsies, serious consequences arise for the patient. Patients will be at greater risk for morbidity and mortality associated with liver biopsies such as postprocedural pain, bleeding, and (less often) pneumothorax. Also, an increase in liver biopsies has socioeconomic consequences such as absence from work. Unnecessary liver biopsies should be avoided, and there is a need for alternative noninvasive and reliable methods to monitor methotrexate-induced liver injury in patients with psoriasis. Several noninvasive methods have been tested as a screening for liver fibrosis and liver cirrhosis (eg, the Fibrotest, Fibroscan, and PIIINP).^{11,13,17,25} Another serious consequence would be the risk of missed pathologic findings that would necessitate discontinuing methotrexate treatment or undergoing another liver biopsy in 6 months.

This study was composed of a rereview of 160 liver biopsy specimens by 1 of us (J.H.v.K.). However, the retrospective nature of the study has some limitations, which might be reflected in the differences in the results found. Slides could have lost some of their stains, and observation of slides serially (by the second pathologist) or in-

dividually (by the first pathologist) could have resulted in some of the differences.

One biopsy specimen was excluded from the study because it was too small to evaluate the degree of fibrosis. A hepatologist experienced in performing liver biopsies and in repeating liver biopsy procedures is essential for obtaining adequate specimens and for the safety of the patient.

Based on this study, we conclude that the interobserver reliability of the Roenigk classification is good and that it can be used as a scoring system for methotrexate-induced liver damage. However, the clinical consequences of rereview were substantial. Experienced pathologists with an interest in liver pathologic conditions are recommended, as well as particular attention to biopsy specimens with Roenigk grades 3a and 3b. The search for noninvasive alternatives to liver biopsy should be continued.

Accepted for Publication: April 6, 2007.

Correspondence: Maartje A. M. Berends, MD, Department of Dermatology, Radboud University Nijmegen Medical Centre, PO Box 9101, 6525 GL Nijmegen, the Netherlands (m.berends@derma.umcn.nl).

Author Contributions: Dr Berends had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. *Study concept and design:* Berends, van Oijen, van de Kerkhof, Drenth, van Krieken, and de Jong. *Acquisition of data:* Berends, Snoek, Drenth, van Krieken, and de Jong. *Analysis and interpretation of data:* Berends, van Oijen, van de Kerkhof, Drenth, van Krieken, and de Jong. *Drafting of the manuscript:* Berends, Snoek, Drenth, van Krieken, and de Jong. *Critical revision of the manuscript for important intellectual content:* Berends, van Oijen, van de Kerkhof, Drenth, van Krieken, and de Jong. *Statistical analysis:* van Oijen and de Jong. *Obtained funding:* van de Kerkhof. *Administrative, technical, or material support:* Berends, Drenth, and de Jong. *Study supervision:* van de Kerkhof, Drenth, van Krieken, and de Jong.

Financial Disclosure: None reported.

REFERENCES

1. Aithal GP, Haugk B, Das S, Card T, Burt AD, Record CO. Monitoring methotrexate-induced hepatic fibrosis in patients with psoriasis: are serial liver biopsies justified? *Aliment Pharmacol Ther.* 2004;19(4):391-399.
2. Boffa MJ, Chalmers RJ, Haboubi NY, Shomaf M, Mitchell DM. Sequential liver

- biopsies during long-term methotrexate treatment for psoriasis: a reappraisal. *Br J Dermatol*. 1995;133(5):774-778.
3. Nyfors A. Liver biopsies from psoriatics related to methotrexate therapy, 3: findings in post-methotrexate liver biopsies from 160 psoriatics. *Acta Pathol Microbiol Scand A*. 1977;85(4):511-518.
 4. Roenigk HH Jr, Maibach HI, Weinstein GP. Methotrexate therapy for psoriasis: guideline revisions. *Arch Dermatol*. 1973;108(1):35.
 5. Zachariae H. Liver biopsies and methotrexate: a time for reconsideration? *J Am Acad Dermatol*. 2000;42(3):531-544.
 6. Zachariae H, Grunnet E, Sogaard H. Liver biopsy in methotrexate-treated psoriatics: a re-evaluation. *Acta Derm Venereol*. 1975;55(4):291-296.
 7. Campbell MS, Reddy KR. Review article: the evolving role of liver biopsy. *Aliment Pharmacol Ther*. 2004;20(3):249-259.
 8. Grismer LE, Gill SA, Harris MD. Liver biopsy in psoriatic arthritis to detect methotrexate hepatotoxicity. *J Clin Rheumatol*. 2001;7(4):224-227.
 9. Roenigk HH Jr, Auerbach R, Maibach H, Weinstein G, Lebwohl M. Methotrexate in psoriasis: consensus conference. *J Am Acad Dermatol*. 1998;38(3):478-485.
 10. Bravo AA, Sheth SG, Chopra S. Liver biopsy. *N Engl J Med*. 2001;344(7):495-500.
 11. Chalmers RJ, Kirby B, Smith A, et al. Replacement of routine liver biopsy by procollagen III aminopeptide for monitoring patients with psoriasis receiving long-term methotrexate: a multicentre audit and health economic analysis. *Br J Dermatol*. 2005;152(3):444-450.
 12. MacDonald A, Burden AD. Noninvasive monitoring for methotrexate hepatotoxicity. *Br J Dermatol*. 2005;152(3):405-408.
 13. Maurice PD, Maddox AJ, Green CA, Tatnall F, Schofield JK, Stott DJ. Monitoring patients on methotrexate: hepatic fibrosis not seen in patients with normal serum assays of aminoterminal peptide of type III procollagen. *Br J Dermatol*. 2005;152(3):451-458.
 14. Roenigk HH, Auerbach R, Weinstein GD. Use of methotrexate in psoriasis. *Arch Dermatol*. 1972;105(3):363-365.
 15. Richard S, Guerret S, Gerard F, Tebib JG, Vignon E. Hepatic fibrosis in rheumatoid arthritis patients treated with methotrexate: application of a new semi-quantitative scoring system. *Rheumatology (Oxford)*. 2000;39(1):50-54.
 16. West SG. Methotrexate hepatotoxicity. *Rheum Dis Clin North Am*. 1997;23(4):883-915.
 17. Zachariae H, Heickendorff L, Søgaard H. The value of amino-terminal propeptide of type III procollagen in routine screening for methotrexate-induced liver fibrosis: a 10-year follow-up. *Br J Dermatol*. 2001;144(1):100-103.
 18. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159-174.
 19. Cicchetti DV, Feinstein AR. High agreement but low κ , II: resolving the paradoxes. *J Clin Epidemiol*. 1990;43(6):551-558.
 20. French METAVIR Cooperative Study Group. Intraobserver and interobserver variations in liver biopsy interpretation in patients with chronic hepatitis C. *Hepatology*. 1994;20(1, pt 1):15-20.
 21. Brunt EM. Grading and staging the histopathological lesions of chronic hepatitis: the Knodell histology activity index and beyond. *Hepatology*. 2000;31(1):241-246.
 22. Goldin RD, Goldin JG, Burt AD, et al. Intra-observer and inter-observer variation in the histopathological assessment of chronic viral hepatitis. *J Hepatol*. 1996;25(5):649-654.
 23. Hunt N, Fleming K. Reproducibility of liver biopsy grading and staging. *Liver*. 1999;19(3):169-170.
 24. Westin J, Lagging LM, Wejstål R, Norkrans G, Dhillon AP. Interobserver study of liver histopathology using the Ishak score in patients with chronic hepatitis C virus infection. *Liver*. 1999;19(3):183-187.
 25. Berends MA, Snoek J, de Jong EM, et al. Biochemical and biophysical assessment of MTX-induced liver fibrosis in psoriasis patients: Fibrotest predicts presence and Fibroscan® predicts the absence of significant liver fibrosis. *Liver Int*. 2007;27(5):639-645.