

Herbal hepatotoxicity and WHO global introspection method

Rolf Teschke, * Axel Eickhoff, * Albrecht Wolff, ** Christian Frenzel, *** Johannes Schulze****

*Department of Internal Medicine II, Division of Gastroenterology and Hepatology, Klinikum Hanau, Academic Teaching Hospital of the Medical Faculty of the Goethe University Frankfurt, Main, Germany.

**Department of Internal Medicine II, Division of Gastroenterology, Hepatology and Infectious Diseases, Friedrich Schiller University, Jena, Germany.

***Department of Medicine I, University Medical Center Hamburg Eppendorf, Germany.

****Office of the Dean, Medical Faculty of the Goethe University Frankfurt/Main, Germany.

ABSTRACT

Herbal hepatotoxicity is a rare but highly disputed disease because numerous confounding variables may complicate accurate causality assessment. Case evaluation is even more difficult when the WHO global introspection method (WHO method) is applied as diagnostic algorithm. This method lacks liver specificity, hepatotoxicity validation, and quantitative items, basic qualifications required for a sound evaluation of hepatotoxicity cases. Consequently, there are no data available for reliability, sensitivity, specificity, positive and negative predictive value. Its scope is also limited by the fact that it cannot discriminate between a positive and a negative causality attribution, thereby stimulating case overdiagnosing and overreporting. The WHO method ignores uncertainties regarding daily dose, temporal association, start, duration, and end of herbal use, time to onset of the adverse reaction, and course of liver values after herb discontinuation. Insufficiently considered or ignored are comedications, preexisting liver diseases, alternative explanations upon clinical assessment, and exclusion of infections by hepatitis A-C, cytomegalovirus (CMV), Epstein-Barr virus (EBV), herpes simplex virus (HSV), and varicella zoster virus (VZV). We clearly prefer as alternative the scale of CIOMS (Council for International Organizations of Medical Sciences) which is structured, quantitative, liver specific, and validated for hepatotoxicity. In conclusion, causality of herbal hepatotoxicity is best assessed by the liver specific CIOMS scale validated for hepatotoxicity rather than the obsolete WHO method that is liver unspecific and not validated for hepatotoxicity. CIOMS based assessments will ensure the correct diagnosis and exclude alternative diagnosis that may require other specific therapies.

Key words. Herb induced liver injury. CIOMS. Kava. Pelargonium sidoides. Herbalife products.

INTRODUCTION

The use of herbal drugs and supplements is popular worldwide and perceived as safe, though some herbal products may cause rare adverse reactions including liver injury.¹ Confirmation of herbal hepatotoxicity represents a particular clinical challenge due to specific disease characteristics and numerous

confounding factors.² Herbs and herbal products contain dozens of various chemicals, rendering compound-specific causality attribution even more complex, for instance, in kava,^{3,4} Herbalife products,⁵⁻⁸ and Greater Celandine.^{9,10} Additionally, the quality of the various causality assessment algorithms is unclear for cases of herb induced liver injury (HILI) as compared to drug induced liver injury (DILI).

Systematic analysis of causality assessment methods is lacking for HILI cases,¹¹ for DILI cases, a study of 2008 reviewed 61 publications of DILI over the last decade from the PubMed database.¹² In 38/61 reports (62.3%), no specific causality assessment method was mentioned; presumably, the evaluation was based on the *ad hoc* approach and thereby on the physicians' judgement lacking any valid criteria for causality assessment. The scale of CIOMS (Council for International Organizations of Medical

Correspondence and reprint request: Rolf Teschke, M.D. Professor of Medicine.

Department of Internal Medicine II, Klinikum Hanau, Academic Teaching Hospital of the Goethe University of Frankfurt, Main

Leimenstrasse 20, D-63450 Hanau, Germany

Tel.: +49-6181/21859

E mail: rolf.teschke@gmx.de

Manuscript received: October 16, 2012.

Manuscript accepted: October 24, 2012.

Sciences) was used in 10 cases (16.4%), the Naranjo scale in 8 cases (13.1%), and the WHO global introspection method (WHO method) in 2 of 61 cases (3.3%).¹² Similarly but not identically, HILI case causality was assessed by the ad hoc approach,¹³⁻¹⁶ the Karch and Lasagna method,⁸ the Naranjo scale,^{17,18} the WHO method,^{5-7,19-21} and the CIOMS scale.^{7,9,20-26} The suitability of these evaluations was subsequently discussed.^{25,27-43}

The *ad-hoc* approach lacks validity criteria for hepatotoxicity cases^{13-16,36} and should therefore be abandoned,³⁶ as should the old and liver unspecific Karch and Lasagna method,⁴⁴ both applied in HILI cases.^{8,13-16} They use subjective judgement for many steps,⁴⁴ making this method prone to bias.⁴⁵ Essential liver specific elements for HILI diagnosis are lacking in the Naranjo scale, which relates toxic drug reactions to pharmacological drug actions rather than specifically to idiosyncratic reactions like hepatotoxicity.⁴⁶ Notably, the Naranjo scale is neither discussed, not even mentioned in reviews and comprehensive surveys of causality assessment methods for liver injury due to drugs, herbs, and dietary supplements,⁴⁷⁻⁵¹ details of its weakness were provided in two review articles^{45,52} and one current opinion article.⁴² This scale lacks specificity for hepatotoxicity by omitting its particular clinical and chronological characteristics;⁵² it is not validated for hepatotoxicity.⁴²

This review focuses on the WHO method,⁵³ which has been applied in suspected herbal hepatotoxicity,^{5-7,19-21,40,41} and discusses its strengths and shortcomings compared to the preferred CIOMS scale to ascertain correct diagnoses and causality levels.^{7,9,20-26,51,54,55}

HERBS, HERBAL DRUGS, AND HERBAL SUPPLEMENTS

Herbs as natural products primarily are used as teas and food additives, whereas manufactured herbal products include drugs and dietary supplements. Herbal drugs are under strict regulatory surveillance, whereas for herbal supplements the regulatory control is less stringent. Most importantly, guidelines exist for quality assurance of pharmaceutical products and monitoring of herbal medicines safety in pharmacovigilance systems. Herbal drugs and herbal supplements should be produced by Good Manufacturing Practices (GMP) including Good Agricultural Practices (GAP). Despite recommendations for quality improvements, batch and product variability including species to species

variation is not unusual.⁴³ Additional concerns may be raised by adulterants, impurities, contaminants, or misidentified herbs.

HERBAL HEPATOTOXICITY

Clinical features of herbal hepatotoxicity or drug induced liver injury are quite similar, idiosyncratic by nature and mostly indistinguishable by laboratory values.^{2,43,51} Based on HILI case series with valid causality results, detailed characteristics of herbal hepatotoxicity as a specific disease have now emerged.^{9,10,23,26,56} Among these are age range, sex ratio, daily dose, treatment duration, latency period, clinical symptoms, comedication, positive reexposure test, laboratory constellation, liver histology, and clinical outcome. Herbal hepatotoxicity is a diagnosis of exclusion; hence, several hundred of other liver diseases with similar characteristics of clinical features, laboratory values, and liver histology have to be differentiated.^{2,12,27,43,45,47-52} To further complicate the situation, 60 different herbs and herbal products have been published as potentially hepatotoxic in 185 publications on herbal hepatotoxicity,¹¹ but adequate causality evaluation was rarely performed.^{11,43}

CAUSALITY ASSESSMENT

Confirming suspected herbal hepatotoxicity is a diagnostic challenge since established laboratory markers to prove a clinical diagnosis are lacking.⁵¹ The primary physician caring for a patient with suspected herbal hepatic injury will start with a clinical assessment to estimate the causal relationship.³⁶ Subsequent evaluation requires stringent diagnostic algorithms to exclude or verify the diagnosis and provide the accurate degree of causality. However, various HILI cases were assessed for herbal products^{5-7,19-21,40,41} by the WHO method⁵³ rather than the CIOMS scale^{51,54} with discrepant results, which raised the question whether these methods are valid.^{20,21,36,40,41,43} Problems with the WHO method are not evident for DILI since only 3.3% of its cases had been submitted to this causality assessment method.¹²

WHO METHOD

The WHO method consists of the WHO scale using broad criteria (Table 1) and a global introspection by experts,⁵³ raising the question to what extent this method may be applicable for assessing

Table 1. WHO scale.

Items of the WHO scale:

1. Certain causality.
 - Event or laboratory test abnormality, with plausible time relationship to drug intake that cannot be explained by disease or other drugs.
 - Response to withdrawal plausible (pharmacologically, pathologically).
 - Event definitive pharmacologically or phenomenologically (i.e. a recognized pharmacological phenomenon or an objective and specific medical disorder).
 - Rechallenge satisfactory, if necessary.
2. Probable causality.
 - Event or laboratory test abnormality, with reasonable time relationship to drug intake.
 - Unlikely to be attributed to disease or other drugs.
 - Response to withdrawal clinically reasonable.
 - Rechallenge not required.
3. Possible causality.
 - Event or laboratory test abnormality, with a reasonable time relationship to drug intake.
 - Could also be explained by disease or other drugs.
 - Information on drug withdrawal may be lacking or unclear.
4. Unlikely causality.
 - Event or laboratory test abnormality, with a time relationship to drug intake that makes a relationship improbable (but not impossible).
 - Disease or other drugs provide plausible explanations.
5. Unclassified causality.
 - Event or laboratory test abnormality.
 - More data for a proper assessment needed, or
 - Additional data under examination.
6. Unassessable causality.
 - Report suggesting an adverse reaction.
 - Cannot be judged because information is insufficient or contradictory.
 - Data cannot be supplemented or verified.

Details are derived from WHO.⁵³

causality in HILI cases with their characteristic features. Analyzing the WHO scale item by item (Table 1), it appears that the various criteria refer to general ADRs and omit liver specific characteristics, rendering the scale disputable.³⁶ In particular, core elements for HILI causality assessment are missing

(Table 2), ignoring current knowledge.^{51,54} Global introspection in itself is an issue of concern.^{39-42,57}

Event

For HILI case assessment, an event is a facultative requirement in the WHO scale (Table 1). The event itself is not further defined but may include symptoms usually reported by HILI patients¹⁰ like weakness, anorexia, nausea, vomiting, abdominal pains, dark urine, acholic stool, itching and jaundice. The criterion of event therefore appear vague.

Laboratory test abnormality

An undefined laboratory test abnormality is another facultative requirement in the WHO scale (Table 1). This implies that even isolated increases of serum γ -glutamyltransferase or a marginally elevated alanine aminotransferase (ALT), aspartate aminotransferase (AST), or alkaline phosphatase (ALP) may suffice as diagnostic criterion to provide a certain, probable or possible causality in HILI cases. Though clear laboratory criteria of hepatotoxicity have been published,^{51,54} respective definitions are lacking in the WHO scale (Tables 1 and 2). Even worse, actual ALT values were missing in 3/15 cases (20%)²⁰ and 6/13 cases (46%)²¹ of assumed HILI by *Pelargonium sidoides* (PS); WHO scale assessment ignored these shortcomings and perpetuated unwarranted causality.^{20,21,58}

In suspected HILI cases, liver enzyme abnormalities do not necessarily origin in the liver since liver involvement unrelated to herbs and drugs is common in numerous primarily nonhepatic diseases.^{37,55} Therefore, liver enzyme abnormalities must be viewed in the context of other parameters and clinical conditions, preventing erroneous diagnoses that are not uncommon in assumed HILI and DILI cases.³⁷

Time relationship

Clinical event or laboratory abnormality are considered valid triggers, provided the time course to herbal intake is plausible or reasonable as presented in the WHO scale (Table 1). These terms, however, remain undefined, plausible or reasonable are open for discussion, the answer must be inaccurate. In HILI cases, time to symptom may range from seven days to one year,^{10,56} leaving a wide time frame for plausible and reasonable. This is crucial to si-

tuations when event or laboratory test abnormality appear days or weeks after cessation of herbal intake. Though temporal association is prerequisite for a causal association, this aspect is poorly handled by the WHO scale, i.e. in cases of suspected HILI by PS.^{20,21,40,41,58} Temporal association between PS use and liver disease could not be confirmed by the CIOMS scale in 8/13 cases (61%),²¹ 5/15 cases (33%),²⁰ and 2/6 cases (33%)⁵⁸ as initially assessed by the WHO method,^{20,21,58} but regulatory decisions still insisted on causality.⁵⁸

Disease

Establishing HILI requires that the symptoms or laboratory tests cannot be explained by disease as mentioned in the WHO scale (Table 1), but suggestions are lacking how to exclude other diseases (Tables 1 and 2). HILI has to be differentiated from multiple liver diseases and liver involvements.⁵¹ Among these are the biliary system including gall bladder, pancreas, small intestine, heart, and endocrine organs. In addition, systemic diseases and

Table 2. Core elements for causality assessment of HILI cases by the WHO scale in comparison with the CIOMS scale.

Details of the individual causality assessment methods	WHO scale	CIOMS scale
• Accurate time frame of latency period.	No	Yes
• Detailed time frame of challenge.	No	Yes
• Clear time frame of dechallenge.	No	Yes
• Recurrent ALT or ALP increase.	No	Yes
• Definition of risk factors.	No	Yes
• Details to exclude alternative diagnoses.	No	Yes
• Assessment of HAV, HBV, HCV.	No	Yes
• Assessment of CMV, EBV, HSV, VZV.	No	Yes
• Liver and biliary tract imaging.	No	Yes
• Color Doppler sonography of liver vessels.	No	Yes
• Assessment of preexisting diseases.	No	Yes
• Evaluation of cardiac hepatopathy.	No	Yes
• Individual score of alternative diagnoses.	No	Yes
• Qualified score of individual comedication.	No	Yes
• Scoring of prior hepatotoxicity by the herb.	No	Yes
• Search for unintended reexposure.	Yes	Yes
• Definition of unintended reexposure.	No	Yes
• Qualified score of unintended reexposure.	No	Yes
• Laboratory criteria for hepatotoxicity.	No	Yes
• Laboratory hepatotoxicity pattern.	No	Yes
• Liver specific method.	No	Yes
• Structured, liver related method.	No	Yes
• Quantitative, liver related method.	No	Yes
• Validated method for hepatotoxicity.	No	Yes

Core elements for assessing causality in cases of herbal hepatotoxicity required for the WHO scale and the CIOMS scale as presented in earlier reports^{36,41,51,53-55}. Latency period indicates time from herb start to symptoms, alternatively to abnormal liver tests. ALT: alanine aminotransferase. ALP: alkaline phosphatase. CIOMS: Council for International Organizations of Medical Sciences. CMV: cytomegalovirus. EBV: Epstein-Barr virus. HAV: hepatitis A virus. HBV: hepatitis B virus. HCV: hepatitis C virus. HSV: herpes simplex virus. VZV: varicella zoster virus.

Table 3. Specific criteria of the WHO scale in comparison with the CIOMS scale used for causality assessment of suspected herbal hepatotoxicity.

Method of causality assessment	Specific criteria of causality assessment methods			
	Structured	Quantitative	Liver-specific	Liver-validated
• WHO scale	Yes	No	No	No
• CIOMS scale	Yes	Yes	Yes	Yes

Compilation of details derived from previous reports.^{11,36,40,41,43,51,53-55} CIOMS scale refers to both the original scale^{54,55} and its update.⁵¹ Liver-specific and liver-validated criteria reflect hepatotoxicity criteria.

other conditions have to be excluded, like general infections, sepsis, rhabdomyolysis, seizures, heat stroke, or polytrauma, to name just a few.

The WHO scale does not require liver and biliary tract imaging or color Doppler sonography of liver vessels to exclude alternative causes (Tables 1 and 2), a problem recognized in assumed HILI.^{20,21} Hence, in 20/28 cases (71%) judged as likely caused by PS by the WHO method, results of abdominal imaging were not described.^{20,21} Also, the WHO scale does not ask specifically for exclusion of virus hepatitis (Tables 1 and 2); again, regulatory assessment of suspected HILI cases with the WHO scale provided causality⁵⁸ while ignoring insufficient exclusion of hepatitis as alternative causes in the study group of 28 patients with assumed HILI by PS.^{20,21} In this cohort, hepatitis A-C was not excluded in 68-71%, as were infections by CMV, EBV, HSV, and VZV in 86-100%.^{20,21} Thus, case assessment by the WHO scale was not reproducible.

Erroneous diagnoses were found in up to 47.1% of all initially suspected hepatotoxicity cases by herbs and drugs,^{9,59} and the culprit remained undetected in up to 38% of severe liver disease.⁶⁰ These data call for a stringent causality assessment in assumed HILI cases to verify or exclude alternative diagnoses, a goal not provided by the WHO scale.

Drugs

To establish HILI as a firm diagnosis, synthetic and other herbal drugs as well as herbal dietary supplements should be excluded as causes for the adverse event; details for the evaluation of comedications are not provided in the WHO scale (Tables 1 and 2). In half of assumed HILI cases, comedication was reported,^{20,21} but this was not sufficiently analyzed and not considered as possible alternative causes by regulatory assessments using the WHO scale.^{20,21,58} Misclassification may result in banning harmless compounds while leaving hepatotoxins in use.

Response to withdrawal

Another key item of the WHO scale is a plausible or reasonable response to drug withdrawal; again, details how to assess these plausible and reasonable features are lacking (Table 1). The time frame of response to withdrawal in suspected HILI cases varies from days to months, and differentiation is necessary between short term improvement and time until complete recovery.^{9,10,26} None of these criteria is considered in the WHO scale (Tables 1 and 2).

For HILI cases, a valid parameter is the fall of ALT or ALP levels during withdrawal,^{51,54} no specific enzyme activity is required in the WHO scale (Table 1). Consequently, little attention is paid to dechallenge characteristics when HILI cases are evaluated by the WHO scale.^{5-7,19-21,58} A detailed analysis revealed that ALT levels during withdrawal were reported in only 15/28 cases (54%), normalization of ALT values was reported in 4/28 cases (14%).^{20,21} Therefore, the WHO scale does not encourage completing laboratory data during the course of dechallenge.

Causality for the herb is highly suggestive if ALT decreases $\geq 50\%$ within 8 days and suggestive if ALT decreases $\geq 50\%$ within 30 days after cessation of intake; if ALT decreases $\geq 50\%$ after 30 days, causality is inconclusive and lacking if ALT decreases $< 50\%$ after 30 days or recurrently increases.^{51,54} No enzyme value contributes specifically in the WHO scale (Tables 1 and 2) or was considered in assessing suspected HILI cases with the WHO scale.^{5-7,19-21,58}

Pharmacological or phenomenological event

For a certain causality, an event has to be observed that is pharmacologically or phenomenologically definitive (i.e. a recognized pharmacological phenomenon or an objective and specific medical disorder) (Table 1). For some ADRs, events are plausible on pharmacological grounds, e.g. a ligand may bind specifically to a receptor. This is, however, difficult to establish for plant-derived products with multiple constituents even in a single herb. It is thus more common that plausibility of HILI is based on phenomenological considerations.

Rechallenge

For a certain causality in the WHO scale, rechallenge is required if necessary, and this test should be satisfactory (Table 1). This implies that a reexposure test is not mandatory, while criteria for a satisfactory test are lacking. For hepatotoxicity, however, some prerequisites are mandatory to ensure transparency and reproducibility of a rechallenge to avoid arbitrary judgements.⁶¹ First, a baseline ALT value below 5N is required after the withdrawal and before the reexposure, with N as the upper limit of normal. Second, during reexposure the ALT value must be at least doubled as compared to the baseline value before reexposure. Both criteria are

obligatory for a positive reexposure test; otherwise, the test is negative. If necessary information is not presented, the test is uninterpretable.⁶¹ Time to onset of symptoms or increased liver values after reexposure should be 1-15 days rather than ≥ 16 days, providing additional strength.^{51,54}

Since specific surrogate markers are lacking, positive reexposure tests are considered as gold standard to prove causality in hepatotoxicity cases.^{51,54,55} However, HILI cases with a positive reexposure test assessed by the CIOMS scale should be reevaluated⁵⁻⁸ whether specific criteria of a positive reexposure test are indeed fulfilled.⁶¹

Confounding variables

Assessment of published HILI cases is commonly impeded by numerous confounding variables.^{5-11,13-26} These include, for example, uncertainty of herbal product quality, poor case data quality, inconsistencies in case data presentation and alternative diagnoses, undisclosed indication, insufficient adverse event definition, lack of temporal association and dechallenge, missing or inadequate evaluation of alcohol use, comedication, comorbidity, and uninterpretable reexposure test.^{25,33} Confounding variables also play a role in suspected HILI cases with causality assessment by the WHO method.^{5-8,19-21,36,40,41,58}

Overdiagnosing and overreporting

A major problem of the WHO method is its tendency of overdiagnosing and overreporting due to ill defined criteria, as shown by other possible primary diagnoses and lack of causality in misattributed HILI, i.e. by PS.^{20,21} For pharmacovigilance purposes, overreporting may result in obscuring problematic compounds; this problem may easily be counteracted by better strategies of the overall assessment approach.²¹

Pharmacovigilance may be improved by three measures: first, improvement in case data quality when presented as spontaneous reports, and early elimination of cases with poor data quality and lack of causality; second, professional case assessment by skilled hepatologists with appropriate clinical evaluation and causality attribution methods; and third, inclusion of cases only when causality for the respective herb has clearly been established by appropriate methods. Emphasis is put on high quality of causality assessment, which may yield fewer but well validated cases, rather than on quantity criteria independent of data quality.

Core elements

The WHO scale lacks a check list with specific core elements characteristic for HILI cases (Table 1), which should be individually assessed (Table 2). Lack of an appropriate check list and the retrospective use of the WHO scale in cases of suspected HILI inevitably creates subjective results of causality assessments^{20,21,40,53,58} that are open for major discussions.^{20,21,36,40,41} Core elements are well defined and listed for DILI^{51,62,54} and HILI.^{42,43,51} For HILI case assessments, all core data elements should prospectively be collected, beginning at the time of the first suspicion of HILI by the treating physician; concomitantly, these data are then to be submitted to further causality assessment by validated methods such as the CIOMS scale. In this context, using the WHO scale is obsolete since this scale prevents rather than promotes valid causality assessment of suspected HILI cases.

Liver specificity

Liver injury has organ-specific properties that are missing in the WHO scale (Tables 1 and 2); consequently, the WHO scale is a liver unspecific causality assessment method and is not applicable to suspected HILI cases.³⁶ This scale ignores differences between organs and fails to consider particular clinical and chronological features of HILI. Among these are a missing definition of liver injury as ADR, lack of clear time frames of latency period, undefined time frame of dechallenge characteristics, no consideration of risk factors, insufficient evaluation of comedication, and lacking definition of a positive reexposure test. This leaves many parameters open for individual interpretation and discussion.

Validation

The WHO scale (Table 1) has not been based on or validated by a gold standard for hepatotoxicity, it is not quantitative and not liver specific (Table 3).^{20,21,36,40,41,53} In particular, data are lacking for reliability, sensitivity, specificity, positive and negative predictive value.^{36,40,41,53} Its scope is also limited by the fact that it cannot discriminate between a positive and a negative causality attribution, thereby stimulating case overdiagnosing and overreporting. Due to these shortcomings, the WHO method is not mentioned or recommended in review articles and textbooks on causality assessment for hepatotoxicity.^{2,24,27,45,47-52} It has been viewed as obsolete for hepatotoxicity.^{36,41}

Table 4. Items of the CIOMS scale required for causality assessment in HILI cases.

Items for hepatocellular injury	Possible Score	Patient's Score	Items for cholestatic (\pm hepatocellular) injury	Possible Score	Patient's Score
1. Time to onset from the beginning of the herb. 5-90 days (rechallenge: 1-15 days). < 5 or > 90 days (rechallenge: > 15 days).	+2 +1		1. Time to onset from the beginning of the herb. 5-90 days (rechallenge: 1-90 days). < 5 or > 90 days (rechallenge: > 90 days).	+2 +1	
2. Time to onset from cessation of the herb. \leq 15 days (except for slowly metabolized herbal chemicals: > 15 days).	+1		2. Time to onset from cessation of the herb. \leq 30 days (except for slowly metabolized herbal chemicals: > 30 days).	+1	
3. Course of ALT after cessation of the herb. • Percentage difference between ALT peak and N. Decrease \geq 50 % within 8 days. Decrease \geq 50 % within 30 days. No information or continued herbal use. Decrease \geq 50 % after the 30th day. Decrease < 50 % after the 30th day or recurrent increase.	+3 +2 0 0 -2		3. Course of ALP after cessation of the herb. • Percentage difference between ALP peak and N. Decrease \geq 50 % within 180 days. Decrease < 50 % within 180 days. No information, persistence, increase, or continued herbal use.	+2 +1 0	
4. Risk factor ethanol. Yes. No.	+1	0	4. Risk factor ethanol or pregnancy. Yes. No.	+1	0
5. Risk factor age. \geq 55 years. < 55 years.	+1 0		5. Risk factor age. \geq 55 years. < 55 years.	+1 0	
6. Concomitant herbs(s) and drug(s). None or no information. Concomitant herb or drug with incompatible time to onset. Concomitant herb or drug with compatible or suggestive time to onset. Concomitant herb or drug known as hepatotoxin and with compatible or suggestive time to onset. Concomitant herb or drug with evidence for is role in this case (positive rechallenge or validated test).	0 0 -1 -2 -3		6. Concomitant herbs(s) and drug(s). None or no information. Concomitant herb or drug with incompatible time to onset. Concomitant herb or drug with compatible or suggestive time to onset. Concomitant herb or drug known as hepatotoxin and with compatible or suggestive time to onset. Concomitant herb or drug with evidence for is role in this case (positive rechallenge or validated test).	0 0 -1 -2 -3	
7. Search for non herb causes. • Group I (6 causes). Anti-HAV-IgM. Anti-HBc-IgM, HBV-DNA.			7. Search for non herb causes. • Group I (6 causes). Anti-HAV-IgM. Anti-HBc-IgM, HBV-DNA.		

Anti-HCV, HCV-RNA.
Hepatobiliary sonography/Colour Doppler sonography of liver vessels/Endosonography.

Alcoholism (AST/ ALT \geq 2).
Acute recent hypotension history (particularly if underlying heart disease).

- Group II (5 causes).
Complications of underlying disease(s).

Infection suggested by PCR and titre change for:

- CMV (Anti-CMV-IgM, Anti-CMV-IgG).
- EBV (Anti-EBV-IgM, Anti-EBV-IgG).
- HSV (Anti-HSV-IgM, Anti-HSV-IgG).
- VZV (Anti-VZV-IgM, Anti-VZV-IgG).

- Evaluation of group I and II.
All causes-groups I and II-reasonably ruled out. +2
The 6 causes of group I ruled out. +1
5 or 4 causes of group I ruled out. 0
Less than 4 causes of group I ruled out. -2
Non herb cause highly probable. -3

- | | |
|--|----|
| 8. Previous information on hepatotoxicity of the herb. | |
| Reaction labelled in the product characteristics. | +2 |
| Reaction published but unlabelled. | +1 |
| Reaction unknown. | 0 |

- | | |
|---|----|
| 9. Response to readministration. | |
| Doubling of ALT with the herb alone, provided ALT below 5N before reexposure. | +3 |
| Doubling of ALT with the herb(s) and drug(s) already given at the time of first reaction. | +1 |
| Increase of ALT but less than N in the same conditions as for the first administration. | -2 |
| Other situations. | 0 |

Total points for patient

Anti-HCV, HCV-RNA.
Hepatobiliary sonography / Colour Doppler sonography of liver vessels/Endosonography.

Alcoholism (AST/ ALT \geq 2).

Acute recent hypotension history (particularly if underlying heart disease).

- Group II (5 causes).
Complications of underlying disease(s).

Infection suggested by PCR and titre change for:

- CMV (Anti-CMV-IgM, Anti-CMV-IgG).
- EBV (Anti-EBV-IgM, Anti-EBV-IgG).
- HSV (Anti-HSV-IgM, Anti-HSV-IgG).
- VZV (Anti-VZV-IgM, Anti-VZV-IgG).

- Evaluation of group I and II.
All causes-groups I and II-reasonably ruled out. +2
The 6 causes of group I ruled out. +1
5 or 4 causes of group I ruled out. 0
Less than 4 causes of group I ruled out. -2
Non herb cause highly probable. -3

- | | |
|--|----|
| 8. Previous information on hepatotoxicity of the herb. | |
| Reaction labelled in the product characteristics. | +2 |
| Reaction published but unlabelled. | +1 |
| Reaction unknown. | 0 |

- | | |
|---|----|
| 9. Response to readministration. | |
| Doubling of ALP with the herb alone, provided ALP below 5N before reexposure. | +3 |
| Doubling of ALP with the herb(s) and drug(s) already given at the time of first reaction. | +1 |
| Increase of ALP but less than N in the same conditions as for the first administration. | -2 |
| Other situations. | 0 |

Total points for patient

Items of the CIOMS scale⁵⁴ are derived from its updated version.⁵¹ In addition, exclusion of hepatitis E by assessment of anti-HEV-IgM, anti-HEV-IgG, and HEV-RNA is mandatory. The above items refer to the hepatocellular type of injury (left scale) and to the cholestatic (\pm hepatocellular) type (right scale). ALP denotes alkaline phosphatase. ALT: alanine aminotransferase. AST: aspartate aminotransferase. CIOMS: Council for International Organizations of Medical Sciences. CMV: cytomegalovirus. EBV: Epstein Barr virus. HAV: hepatitis A virus. HBc: hepatitis B core. HBV: hepatitis B virus. HCV: hepatitis C virus. HILI: herb induced liver injury. HSV: herpes simplex virus. N: upper limit of the normal range. VZV: varicella zoster virus. Total points and causality: \leq 0: excluded. 1-2: unlikely. 3-5: possible. 6-8: probable. $>$ 8: highly probable.

Global introspection

In assessing the likelihood of drug causality in individual cases of general adverse drug reactions (ADRs), global introspection represents a popular strategy.^{57,63} Specifically, the assessor considers factors that might contribute to a causal link between one or more drugs and an observed ADR; lists these factors, weighs their importance, and decides the probability of drug causation;⁵⁷ no check list or level of strength is given. Not surprisingly, this method lacks validation for any type of ADRs.⁴¹ Both the criteria and the parameters are ambiguous.³⁶ Already 40 years ago, global introspection was criticized as subjective and imprecise since it is based only on expert clinical judgement.⁶³⁻⁶⁵ In 1986, global introspection by experts has been shown to be neither reproducible, nor valid, nor accountable.⁵⁷ Overall, causality levels may be doubtful if based on the WHO method, and assessments should be repeated using a valid method.

CIOMS SCALE

To improve strength of causality assessment in HILI cases, the problems associated with the WHO method have to be addressed. Currently, the method of choice is the CIOMS scale (Table 4) in its original form^{54,55} or its update.⁵¹ The CIOMS scale is structured, quantitative, liver specific, and validated for hepatotoxicity (Table 3); it considers all core elements of hepatotoxicity cases (Tables 2 and 4).^{51,54} In addition, CIOMS based assessment has shown good sensitivity (86%), specificity (89%), positive predictive value (93%), and negative predictive value (78%).⁵⁵ The CIOMS scale was developed by an international expert panel and based on cases with positive reexposure tests as gold standard.^{54,55} Validated reexposure tests meeting the specific criteria are included into the CIOMS scale (Table 4).^{51,54} Prerequisite for the assumption of a relevant liver disease is a value for ALT or ALP of at least $> 2N$, and laboratory evaluation differentiates between a hepatocellular and a cholestatic (\pm hepatocellular) type of injury to choose the correct CIOMS scale for evaluation (Table 4).⁵¹ Therefore, key elements for assessing causality of herbal hepatotoxicity are fulfilled by the CIOMS scale.

The CIOMS scale has successfully been applied in various reports of hepatotoxicity in epidemiological studies, clinical trials, case reports, case series,

regulatory analyses, and genotyping studies.³⁶ The scale provides a range of causality gradings for the responsible agent(s) and clearly delineates liver specific criteria for challenge, dechallenge, exclusion of other and unrelated diseases, comedicated synthetic drugs, herbal drugs, and dietary supplements including herbal ones.^{36,51,54,55}

WHO METHOD AS COMPARED TO CIOMS SCALE

In contrast to other views,⁴⁰ we clearly prefer the CIOMS scale over the WHO method in assessing hepatotoxicity causality, thereby providing the correct diagnosis and excluding alternative diseases that may require other specific therapies.^{11,20,21,36,41} Virtually no single essential element for valid causality assessment as detailed by the CIOMS scale (Tables 2-4) is represented by the WHO method (Tables 1-3); case overreporting by incorrect assessment is prevented using the CIOMS scale.^{20,21,41} The claim of higher sensitivity of the WHO scale compared to the CIOMS scale is difficult to reconcile,^{40,41} since the CIOMS scale has a sensitivity of 86%,⁵⁵ whereas no sensitivity has been published for the WHO method.^{36,40,41,53} Thus, the CIOMS scale but not the WHO method is exclusively designed for hepatotoxicity cases.

A primary care physician suspecting herbal hepatotoxicity can easily use the CIOMS scale. This usability contrasts to the WHO method that obligatorily requires a team of experts, but lacks related quality standards. Results by the CIOMS scale are available within a few minutes at the bedside, whereas those by the obsolete WHO method are presented at best months after the initial suspicion, at a time when decisions have been made long before. Though physician experts are members of the drug commission,⁴⁰ none of these realized flaws like retracted cases, case duplications, and/or alternative diagnoses.^{20,21,41}

CONCLUDING REMARKS

In cases of suspected herbal hepatotoxicity, a sophisticated management of causality assessment is mandatory to ensure the correct diagnosis and to exclude alternative diseases that may require other specific therapies. This goal is achieved by methods that are liver specific and validated for hepatotoxicity, criteria fulfilled by the preferred CIOMS scale but not by the WHO method.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCES

- Colalto C. Unpredictable adverse reactions to herbal products. *J Drug Metab Toxicol* 2012; 3: 2. Available at: <http://www.omicsonline.org/2157-7609/2157-7609-3-e105.pdf> [Accessed October 15, 2012].
- Navarro VJ. Herbal and dietary supplement hepatotoxicity. *Semin Liver Dis* 2009; 29: 373-82.
- Teschke R, Lebot V. Proposal for a Kava Quality Standardization Code. *Food Chem Toxicol* 2011; 49: 2503-16.
- Teschke R, Qiu SX, Xuan TD, Lebot V. Kava and kava hepatotoxicity: requirements for novel experimental, ethnobotanical, and clinical studies based on a review of the evidence. *Phytother Res* 2011; 25: 1262-74.
- Elinav E, Pinkser G, Safadi R, Pappo O, Bromberg M, Anis E, Keinan-Boker L, et al. Association between consumption of Herbalife® nutritional supplements and acute hepatotoxicity. *J Hepatol* 2007; 47: 514-20.
- Schoepfer AM, Engel A, Fattinger K, Marbet UA, Criblez D, Reichen J, Zimmermann A, et al. Herbal does not mean innocuous: ten cases of severe hepatotoxicity associated with dietary supplements from Herbalife products. *J Hepatol* 2007; 47: 521-6.
- Jóhannsson M, Ormarsdóttir S, Ólafsson S. Hepatotoxicity associated with the use of Herbalife. *Laeknabladid* 2010; 96: 167-72.
- Manso G, López-Rivas L, Salgueiro ME, Duque JM, Jimeno FJ, Andrade RJ, Lucena MI. Continuous reporting of new cases in Spain supports the relationship between Herbalife products and liver injury. *Pharmacoepidemiol Drug Saf* 2011; 20: 1080-7.
- Teschke R, Glass X, Schulze J. Herbal hepatotoxicity by Greater Celandine (*Chelidonium majus*): Causality assessment of 22 spontaneous reports. *Regul Toxicol Pharmacol* 2011; 61: 282-91.
- Teschke R, Frenzel C, Glass X, Schulze J, Eickhoff A. Greater Celandine hepatotoxicity: A clinical review. *Ann Hepatol* 2012; 11: 838-48.
- Teschke R, Wolff A, Frenzel C, Schulze J, Eickhoff A. Herbal hepatotoxicity: A tabular compilation of reported cases. *Liver Int* 2012. In press. DOI: 10.1111/j.1478-3231.2012.02864x.
- Tajiri K, Shimizu Y. Practical guideline for diagnosis and early management of drug-induced liver injury. *World J Gastroenterol* 2008; 14: 6774-85.
- BfArM (Bundesinstitut für Arzneimittel und Medizinprodukte, Bonn. Federal Institute for Drugs and Medicinal Products in Germany). Rejection of Drug Risks, Step II. As related to: Kava-Kava (Piper methysticum)-containing, and kavain-containing drugs, including homeopathic preparations with a final concentration up to, and including D4. June 14, 2002. Available at: http://www.spc.int/cis/documents/02_0714_BfArM_Kava_Removal.pdf [Accessed October 15, 2012].
- Benninger J, Schneider HT, Schuppan D, Kirchner T, Hahn EG. Acute hepatitis induced by Greater Celandine (*Chelidonium majus*). *Gastroenterology* 1999; 117: 1234-7.
- Dara L, Hewett J, Lim JK. Hydroxycut hepatotoxicity: A case series and review of liver toxicity from herbal weight loss supplements. *World J Gastroenterol* 2008; 14: 6999-7004.
- Inoue H, Yamazaki S, Shimizu M, Uozki H, Goto T, Ohnishi S, Koike K. Liver injury induced by the Japanese herbal drug kamishoyosan. *Gastroenterol Hepatol* 2011; 7: 692-5.
- Mahady GB, Low Dog T, Barrett ML, Chavez ML, Gardiner P, Ko R, Marles RJ, et al. United States Pharmacopeia review of the black cohosh case reports of hepatotoxicity. *Menopause* 2008; 15: 628-38.
- Sarma DN, Barrett ML, Chavez ML, Gardiner P, Ko R, Mahady GB, Marles RJ, et al. Safety of green tea extract: a systematic review by the US Pharmacopeia. *Drug Saf* 2008; 31: 469-84.
- WHO (World Health Organization). Assessments of the risk of hepatotoxicity with kava products. Geneva, Switzerland: WHO Document Production Services; 2007.
- Teschke R, Frenzel C, Schulze J, Eickhoff A. Spontaneous reports of primarily suspected herbal hepatotoxicity by Pelargonium sidoides: Was causality adequately ascertained? *Regul Toxicol Pharmacol* 2012; 63: 1-9.
- Teschke R, Frenzel C, Wolff A, Herzog J, Glass X, Schulze J, Eickhoff A. Initially purported hepatotoxicity by Pelargonium sidoides: the dilemma of pharmacovigilance and proposals for improvements. *Ann Hepatol* 2012; 11: 500-12.
- EMA: Assessment of case reports connected to herbal medicinal products containing cimicifugae racemosae rhizoma (black cohosh, root). Issued May 8, 2007. Available at: http://www.ema.europa.eu/docs/en_GB/document_library/Herbal_-HMPA_assessment_report/2010/02/WC500074167.pdf [Accessed October 15, 2012].
- Teschke R, Schwarzenboeck A, Hennermann KH. Kava hepatotoxicity: a clinical survey and critical analysis of 26 suspected cases. *Eur J Gastroenterol Hepatol* 2008; 20: 1182-93.
- García-Cortés M, Borraz Y, Lucena MI, Peláez G, Salmerón J, Diago M, Martínez-Sierra MC, et al. Liver injury induced by "natural remedies": an analysis of cases submitted to the Spanish Liver Toxicity Registry. *Rev Esp Enferm Dig* 2008; 100: 688-95 (Article in Spanish).
- Teschke R, Schmidt-Taenzer W, Wolff A. Spontaneous reports of assumed herbal hepatotoxicity by black cohosh: Is the liver unspecific Naranjo scale precise enough to ascertain causality? *Pharmacoepidemiol Drug Saf* 2011; 20: 567-82.
- Teschke R, Glass X, Schulze J, Eickhoff A. Suspected Greater Celandine hepatotoxicity: Liver specific causality evaluation of published case reports from Europe. *Eur J Gastroenterol Hepatol* 2012; 24: 270-80.
- Liss G, Lewis JH. Drug-induced liver injury: what was new in 2008? *Expert Opin Drug Metab Toxicol* 2009; 5: 843-60.
- Teschke R. Letter to the Editor: United States Pharmacopeia review of the black cohosh case reports of hepatotoxicity. *Menopause* 2009; 16: 214.
- Mahady G, Low Dog T, Sarma DN. Letter to the Editor (Reply). *Menopause* 2009; 16: 214-5.
- Teschke R, Bahre R, Genthner A, Fuchs J, Schmidt-Taenzer W, Wolff A. Suspected black cohosh hepatotoxicity-Challenges and pitfalls of causality assessment. *Maturitas* 2009; 63: 302-14.
- Mahady GB, Low Dog T, Sarma DN, Giancaspro GI. Letter to the Editor. Suspected black cohosh hepatotoxicity-Causality assessment versus safety signal. *Maturitas* 2009; 64: 139-40.
- Teschke R, Bahre R, Genthner A, Fuchs J, Schmidt-Taenzer W, Wolf A. Suspected black cohosh hepatotoxicity-Causality assessment versus safety signal. Quality versus quantity. *Maturitas* 2009; 64: 141-2.

33. Teschke R. Black cohosh and suspected hepatotoxicity-inconsistencies, confounding variables, and prospective use of a diagnostic causality algorithm: A critical review. *Menopause* 2010; 17: 426-40.
34. Mahady G, Low Dog T, Sarma ND, Giancaspro GI, Griffiths J. Letter to the Editor. *Menopause* 2010; 17: 1088-9.
35. Teschke R. Letter to the editor: In response. *Menopause* 2010; 17: 1089.
36. Teschke R, Wolff A. Regulatory causality evaluation methods applied in kava hepatotoxicity: Are they appropriate? *Regul Toxicol Pharmacol* 2011; 59: 1-7.
37. Teschke R, Schwarzenboeck A, Schmidt-Taenzer W, Wolff A, Hennermann KH. Herb induced liver injury presumably caused by black cohosh: A survey of initially purported cases and herbal quality specifications. *Ann Hepatol* 2011; 11: 249-59.
38. Mahady G, Dog T, Sarma ND, Griffiths J, Giancaspro G. Letter to the Editor. *Pharmacoepidemiol Drug Saf* 2012; 21: 339-40.
39. Teschke R, Schmidt-Taenzer W, Wolff A. USP suspected herbal hepatotoxicity: Quality of causality assessment is more important than quantity of counted cases, not vice versa. *Pharmacoepidemiol Drug Saf* 2012; 21: 336-38.
40. Stammschulte T, Gundert-Remy U. Letter to the Editor. Spontaneous reports of primarily suspected herbal hepatotoxicity by Pelargonium sidoides: Was causality adequately ascertained? *Regul Toxicol Pharmacol* 2012. In press. Available at: <http://dx.doi.org/10.101016/j.yrtph.2012.06.011> [Accessed October 15, 2012].
41. Teschke R, Frenzel C, Schulze J, Eickhoff A. Suspected herbal hepatotoxicity: the pharmacovigilance dilemma with disputed and obsolete evaluation methods. *Regul Toxicol Pharmacol* 2012. In press. Available at: <http://dx.doi.org/10.1016/j.yrtph.2012.06.010> [Accessed October 15, 2012].
42. Teschke R, Schulze J. Suspected herbal hepatotoxicity: Requirements for appropriate causality assessment by the US Pharmacopeia. *Drug Saf* 2012. In press. DOI: 10.2165.11631960-000000000-00000.
43. Teschke R, Frenzel C, Glass X, Schulze J, Eickhoff A. Herbal hepatotoxicity: A critical review. *Br J Clin Pharmacol*. In press. DOI: 10.1111/j.1365-2125.2012.04395.x
44. Karch FE, Lasagna L. Toward the operational identification of adverse drug reactions. *Clin Pharmacol Ther* 1977; 21: 247-54.
45. García-Cortés M, Stephens C, Lucena MI, Fernández-Castañer A, Andrade RJ, on behalf of the Spanish Group for the Study of Drug-Induced Liver Disease. Causality assessment methods in drug induced liver injury: Strengths and weaknesses. *J Hepatol* 2011; 55: 683-91.
46. Naranjo CA, Busto U, Sellers EM, Sandor P, Ruiz I, Roberts EA, Janecek E, et al. A method for estimating the probability of adverse drug reactions. *Clin Pharmacol Ther* 1981; 30: 239-45.
47. Zimmerman HJ. *Hepatotoxicity*. Philadelphia: Lippincott Williams & Wilkins; 1999.
48. Lee WM. Drug-induced hepatotoxicity. *N Engl J Med* 2003; 349: 474-85.
49. Lee WM, Senior JR. Recognizing drug-induced liver injury: Current problems, possible solutions. *Toxicol Pathol* 2005; 33: 155-64.
50. Watkins PB, Seeff LB. Drug-induced liver injury: Summary of a single topic clinical research conference. *Hepatology* 2006; 43: 618-31.
51. Teschke R, Schwarzenboeck A, Hennermann KH. Causality assessment in hepatotoxicity by drugs and dietary supplements. *Br J Clin Pharmacol* 2008; 66: 758-66.
52. Andrade RJ, Camargo R, Lucena MI, González-Grande R. Causality assessment in drug-induced hepatotoxicity. *Expert Opin Drug Saf* 2004; 3: 329-44.
53. WHO, World Health Organization. The use of the WHO-UMC system for standardised case causality assessment. WHO Collaborating Centre for International Drug Monitoring (Uppsala Monitoring Centre, UMC), Database 2000. Available at: <http://who-umc.org/Graphics/24734.pdf> [Accessed October 15, 2012].
54. Danan G, Bénichou C. Causality assessment of adverse reactions to drugs-I. A novel method based on the conclusions of international consensus meetings: application to drug-induced liver injuries. *J Clin Epidemiol* 1993; 46: 1323-30.
55. Bénichou C, Danan G, Flahault A. Causality assessment of adverse reactions to drugs-II. An original model for validation of drug causality assessment methods: case reports with positive rechallenge. *J Clin Epidemiol* 1993; 46: 1331-6.
56. Teschke R. Kava hepatotoxicity-a clinical review. *Ann Hepatol* 2010; 9: 251-65.
57. Kramer MS. Assessing causality of adverse drug reactions: Global introspection and its limitations. *Drug Inf J* 1986; 20: 433-7.
58. Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM). Pelargoniumhaltige Arzneimittel: Risiko von Leberschädigungen, 28.6. 2012. Available at: <http://www.bfarm.de> [Accessed October 15, 2012].
59. Aithal GP, Rawlins MD, Day CP. Accuracy of hepatic adverse drug reaction reporting in one English health region. *Br Med J* 1999; 319: 154.
60. Bernal W, Auzinger G, Dhawan A, Wendon J. Acute liver failure. *Lancet* 2010; 376: 190-201.
61. Bénichou C. Criteria of drug-induced liver disorders. Report of an international consensus meeting. *J Hepatol* 1990; 11: 272-6.
62. Agarwal VK, McHutchison JG, Hoofnagle JH, Drug-Induced Liver Injury Network (DILIN). Important elements for the diagnosis of drug-induced liver injury. *Clin Gastroenterol Hepatol* 2010; 8: 463-70.
63. Macedo AF, Marques FB, Ribeiro CF, Teixeira F. Causality assessment of adverse drug reactions: comparison of the results obtained from published decisional algorithms and from the evaluations of an expert panel, according to different levels of imputability. *J Clin Pharm Ther* 2003; 28: 137-43.
64. Karch FE, Smith CL, Kerzner B, Mazullo JM, Weintraub M, Lasagna L. Adverse drug reactions a matter of opinion. *Clin Pharmacol Ther* 1976; 19: 489-92.
65. Blanc S, Levenberger P, Berger JP, Brooke EM, Schelling JL. Judgements of trained observers on adverse drug reactions. *Clin Pharmacol Ther* 1979; 25: 493-8.