Regulatory Toxicology and Pharmacology 60 (2011) 290–295

Contents lists available at ScienceDirect

Regulatory Toxicology and Pharmacology

journal homepage: www.elsevier.com/locate/yrtph

What are the best reference values for a normal serum alanine transaminase activity (ALT)? Impact on the presumed prevalence of drug induced liver injury (DILI)

Helmi M'Kada [a,e,1], Mona Munteanu [b,1], Hugo Perazzo [a,1], Yen Ngo [b,1], Nittia Ramanujam [a,1], Françoise Imbert-Bismut [a,e,1], Vlad Ratziu [a,1], Dominique Bonnefont-Rousselot [a,e,f,1], Bernard Souberbielle [c,1,2], Ina Schuppe-Koistinen [d,1], Thierry Poynard [a,*,1] and the DILI group of the SAFE-T consortium

[a] Assistance Publique Hôpitaux de Paris UPMC Liver Center, Paris, France
[b] Biopredictive, Clinical Research, Paris, France
[c] Pfizer, Clinical Research, Ramsgate Road, Sandwich, Kent CT13 9NJ, UK
[d] Astra Zeneca, Södertälje, Sweden
[e] Service de Biochimie Métabolique, Groupe Hospitalier Pitié-Salpêtrière (AP-HP), Paris, France
[f] EA 4466, Faculté des Sciences Pharmaceutiques et Biologiques, Université Paris Descartes, Paris, France

ARTICLE INFO

Article history:
Received 4 January 2011
Available online 24 April 2011

Keywords:
DILI
Biomarkers
ALT
Limit of the normal
Hy's law
Temple criteria
Liver necrosis
Reference values

ABSTRACT

Background: In clinical research, the definition of the upper limit of normal (ULN) is rarely detailed. For alanine transaminase (ALT), there are several definitions of ULN-ALT but no recognized global reference. Furthermore the inter-laboratory variability of results expressed using ULN-ALT is higher than using the actual value of ULN expressed in IU/L. Regulatory agencies still use ULN-ALT for the definition of drug adverse events such as drug induced liver disease (DILI).
Methods: We applied two extreme definitions of ULN-ALT (26 and 66 IU/L) in two populations with different liver disease risk: 7463 consecutive volunteers representative a low risk population, and 6865 consecutive patients hospitalized in a tertiary referral center. The same assay technique was used for both populations on fresh plasma in the same laboratory.
Results: In the low risk population the liver disease estimates ranged from 0% to 1.99% according to ULN-ALT definition and gender; prevalence of liver disease as defined by Temple's criteria (3×ULN) decreased significantly with increased ULN-ALT threshold and prevalence of liver disease was lower in females compared to males (all $P < 0.001$). In the high risk population the estimates of liver disease prevalence ranged from 0.78% to 15.85%; disease prevalence using both Temple's corollary and Hy's law criteria (3×ULN-ALT and bilirubin >34 μmol/L) decreased significantly with increased ULN-ALT threshold and females compared to males. In the low risk population the two major factors associated with ULN variability were gender and BMI.
Conclusion: Artificial statistical modifications of the procedures chosen for the ULN-ALT definition change dramatically the prevalence of DILI estimates. A consensus in liver disease definitions seems mandatory for DILI studies in order to prevent misleading conclusions.

1. Introduction

Serum alanine transaminase (ALT) activity is widely used as the primary reference estimate for identifying drug induced liver disease (DILI) (Ozer et al., 2010). Unfortunately, ALT assays have inter-laboratory variability because of variability in analytical methods but also because of variability in the values for the upper limit of normal (ULN) (Ozer et al., 2010; Piton et al., 1998; Dutta et al., 2009; Myara et al., 2004; Ferard et al., 2006; Imbert-Bismut et al., 2004; Halfon et al., 2002). The latter is mostly related to the

* Corresponding author. Address: Hepatology, Groupe Hospitalier Pitié Salpêtrière, 47 Bd Hôpital, 75013 Paris, France. Fax: +33 1 42 16 14 27.
E-mail address: tpoynard@teaser.fr (T. Poynard).

[1] For the DILI-GHPS Group and the SAFE-T Consortium (SAFE-T IMI Consortium members for DILI workpackage are: BS, ISK, TP, HP, NR).

[2] The personal views expressed in this article may not be understood nor quoted as being made on behalf of or reflecting the position of the IMI or its membership companies.

0273-2300/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.yrtph.2011.04.002

difference in reference populations from which these ULN are calculated (Dutta et al., 2009).

We observed that this ALT inter-laboratory variability was worse when ULN was used instead of using the actual value of ULN expressed in IU/L (Imbert-Bismut et al., 2004; Halfon et al., 2002). This is particularly important clinically, because numerous medical guidelines make reference to ALT expressed as multiples of the ULN (Ozer et al., 2010; Dutta et al., 2009), variations in the definition of normal may have important practical consequences.

The change in methods successively used in Groupe Hospitalier Pitié-Salpêtrière, Paris, France, is an example of how the laboratories generate the ULN. Between 1993 and 1996, there was a common threshold for ULN of 45 IU/L used both for males and females. This threshold corresponded to the mean + 2 standard deviation (SD) of a control population given by the manufacturer and after exclusion of the 5% extreme values. There was no scientific publication describing this control population. Since January 1996, the ULN was based on a study of 2200 apparently healthy blood donors negative for human immunodeficiency virus, hepatitis B virus, and hepatitis C virus markers. It included 1171 men and 880 women. The thresholds were 26 IU/L in women and 35 in men. They were determined by the mean +1 SD after exclusion of the 5% extreme values.

With the increase of patients with Hepatitis C we rapidly observed major inter-laboratories discordances between repeated ALT. ALT assays have inter-laboratory variability because of variability in analytical methods but also because of variability in the values of (Ozer et al., 2010; Piton et al., 1998; Dutta et al., 2009; Myara et al., 2004; Ferard et al., 2006; Imbert-Bismut et al., 2004; Halfon et al., 2002). The latter is mostly related to the difference in reference populations from which these ULN are calculated including the rules of ULN definitions (Dutta et al., 2009). After comparing eight definitions of ULN and the inter-laboratories-variability of nine labs we decided not to expressed ALT values using ULN as a unit, but simply in IU/L (Piton et al., 1998; Halfon et al., 2002).

The increase in the number of biopsies permitted to revisit the performances of ALT as a first line liver test for the main liver injuries using biopsy as an (imperfect) gold standard and receiver operating characteristics curve as the reference method. Using this method no definition of normal value is predetermined, the threshold being chosen according to the goal of the clinician, i.e. favoring negative or positive predictive values.

For the diagnosis of fibrosis stages, we and others rapidly observed that the accuracy of ALT was significantly higher than random but weak in comparison with specifically designed biomarkers such as FibroTest and FibroScan (Castera and Pinzani, 2010; Poynard, 2011).

For the diagnosis of necro-inflammatory grades in patients with chronic viral hepatitis, the accuracy of ALT was high. To our knowledge only one specific biomarker of necrosis activity (ActiTest) has demonstrated a significant higher accuracy in patients with chronic hepatitis C (Poynard et al., 2010).

For the diagnosis of steatosis, the accuracy of ALT was significantly better than random but lower than the accuracy of a specific biomarkers of steatosis (SteatoTest) (Poynard et al., 2005)."

For acute liver disease such as DILI there is no scientific answer because no validation of ALT or of a specific DILI biomarker has been published. Several reasons explained the difficulty of such validation. Severe DILI necrosis with jaundice is a too rare event, less than 1 out of 10,000 exposed patients. Even for a less severe DILI, a validation on a large population is still difficult due to the limitations of biopsy. Therefore ALT "3 times the ULN in the absence of other cause", despite an absence of validation, is still the standard for DILI definition (Ozer et al., 2010).

For DILI studies the definition of abnormal serum ALT activity has 2 important consequences. One relates to the variability of the "Temple's corollary criteria", the first estimate used for the suspicion of necrosis and inflammation, the usual ALT cutoff being 3×ULN (Ozer et al., 2010). The second circumstance, which is related to the estimate used as a marker of severe liver necrosis, is the Hy's criteria (or Hy's law) usually defined as 3×ULN-ALT and total bilirubin >34 μmol/L (i.e. 2×ULN), in the absence of other etiology to explain rise in ALT or bilirubin (Ozer et al., 2010).

The specific aim of the present study was to assess the impact of the variability of the ULN-ALT definition on the prevalence of liver disease as defined by Temple's criteria and Hy's law, two commonly-used estimates of liver necrosis and inflammation to assess DILI. We have used the recommended criteria published by Dutta et al. to assess the impact of changes in the reference population characteristics on ULN threshold (Dutta et al., 2009).

2. Methods

2.1. Subjects

Two populations were included in this study: one population deemed at "low risk of ALT increase" (i.e. general population) and the other group, deemed at "high risk of ALT increase" (i.e. patients referred to a tertiary referral center). The general population included 7463 consecutive apparently healthy volunteers, over 40 years of age, representative of the French general population, who were seen for a free screening program in two French Social Security health examination centers (median 58 years, 45% declaring prescription drugs or OTC drugs). This volunteer population was part of an already published epidemiological study estimating the prevalence of non-overt chronic liver disease using non invasive *biomarkers* (Poynard et al., 2010). *The "high risk of ALT increase" population included 6865 consecutive* patients hospitalized in a tertiary referral center (Groupe Hospitalier Pitié Salpêtrière, Paris, France). The median age of this patient population was 53 years, and 9.9% were hospitalized in the Hepatology department, mainly patients with cirrhosis due to chronic hepatitis C, B, and alcoholic liver disease. None of them had a severe DILI. All medical and surgical specialties were represented in this 1800 bed university hospital. In these two populations the diagnosis of liver injuries was a global estimate. According to the practice in France, biopsy was only performed in case of discordance between non-invasive biomarkers and if accepted by patients.

2.2. Biochemical analysis

The same assay techniques were used for both populations in the same laboratory. Biochemical assays were performed with fresh plasma decanted and stored for a maximum of 72 h at +2–8 °C, under no-light conditions. For ALT, activity measurement used the reference method defined by the International Federation of Clinical Chemistry (IFCC) with pyridoxal phosphate and was calibrated (Ferard et al., 2006; Imbert-Bismut et al., 2004). Total bilirubin was assayed by the diazo-reaction method.

2.3. Statistical methods

We applied the two extreme definitions of ULN-ALT (26 and 66 IU/L respectively) out of 7 published definitions (Piton et al., 1998) to calculate the prevalence of two standard estimates of liver necrosis (Temple's criteria and Hy's law (Ozer et al., 2010). As the normal mean value of ALT is higher in males than in females, the prevalence was also calculated according to gender (Piton et al., 1998). The threshold of ULN ALT 26 IU/L was the mean ALT + 1 standard deviation (SD) after exclusion of the 5% extreme values assessed among 880 healthy female blood donors. The threshold

of ULN-ALT of 66 IU/L was the 95th percentiles assessed in 1033 males blood donors with BMI greater than 23 kg/m^2.

The impact of the following recommended criteria for a reference population were also assessed retrospectively in the "low risk of ALT increase" low risk population: healthy subjects with normal body weight (defined as BMI not greater than 25 kg/m^2), no underlying acute or chronic illnesses, no significant alcohol consumption (defined as no more than 10 g/day for female and 20 g/day for male), and no intake of prescription medicines. We were able to exclude subjects with alcohol consumption, with elevated BMI, with chronic diseases and declared intake of drugs. However, we were not able to eliminate formally intake of nonprescription medicines, such as herbal compounds or dietary supplements (Castera and Pinzani, 2010).

Comparisons between prevalence values were performed using the Fisher exact test and Number Cruncher statistical systems software (NCSS, 2007).

3. Results

The main characteristics of subjects included in the two populations are given in table 1. Prevalences of necro-inflammatory estimates, according to ULN-ALT definition and gender in the two populations, are given in Table 2.

Using the 26 IU/L threshold, the 3-fold ULN-ALT limit was 78 IU/L and using 66 IU/L threshold, the 3-fold ULN-ALT limit was 198 IU/L, with dramatic differences in the distribution of patients fulfilling the Temple's or Hy's law criteria (Fig. 1).

There were highly significant differences (all $P < 0.0001$) between prevalences of liver disease as defined by Temple's criteria according to the different definitions of ULN-ALT, in both the low risk population and the patient population, both for males and females (Table 2).

No increase (0%) was observed in females in the low risk population versus 0.69% using a 26 IU/L threshold for ULN-ALT

Severe estimates (Hy's law) were not observed in the low risk population. In the tertiary center population, however, the prevalence for Hy's law cases varied from 0.78% to 4.38% according to ULN-ALT definition and gender.

Within our low risk population, ULN-ALT was calculated after stratification by various reference populations defined according to recommended criteria, and these values are given in Table 3 (Dutta et al., 2009). The two major factors associated with ULN variability are gender and BMI. In our low risk population the ULN varied from 38 IU/L (95%CI 34–47 IU/L) in female with normal BMI, no medication use, and alcohol consumption not greater than 10 gr/day to 61 IU/L (59–61 IU/L) in male not selected for BMI and alcohol consumption (Table 3).

Table 1
Characteristics of patients included.

Characteristic	Low risk population	Tertiary center
Total number of patients	7463	6865
Male n, (%)	4113 (55.1%)	3900 (57.1%)
Age (years) median, (95% CI)	57.8 (57.6–58.0)	53.0 (52.5–53.7)
Medication use n, (%)	3350 (44.9%)	6865 (100%)
Overt liver disease n, (%)	0 (0%)	678 (9.9%)
ALT (IU/L) median, (95% CI)	23 (23–23)	29 (28–29)
ALT (IU/L) mean, (95% CI)	27 (26–28)	62 (58–67)
Male	31 (30–32)	73 (66–81)
Female	22 (22–23)	48 (43–52)
Total bilirubin (micromol/L) median, (95% CI)	10 (10–10)	7 (7–7)
Total bilirubin (micromol/L) mean, (95% CI)	12 (11.9–12.1)	17 (16–19)

4. Discussion

Given that ALT is one of the most important and most widely ordered laboratory tests, it is unacceptable to have such a wide variability in its ULN, regardless of the origin of the problem (Piton et al., 1998; Dutta et al., 2009). Until a single analytical ALT enzymatic method is used by all and until references values are validated and recognized by all stakeholders (industry, academia and regulatory agencies), we believe that ALT expression should only be reported as actual numerical value expressed in IU/L (Ozer et al., 2010; Imbert-Bismut et al., 2004). The implications for forthcoming DILI studies is when actual numerical value of ALT is compared to ULN, then it is imperative to detail the methods and the reference population used for the ULN definition.

4.1. ALT as a reference of liver disease

ALT is a sensitive but not specific marker for liver injury. Monitoring liver chemistries in drug development studies is key among liver safety measures and the incidence of relatively small elevations in ALT are considered as estimates of concern (http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/ Guidances/UCM174090.pdf, 0000; Weil et al., 2008; Fontana et al., 2010; Kim et al., 2004; Watkins et al., 2008).

In the present study we estimated the very significant impact of the ULN definition variability on standard estimates used in DILI studies, both for detection sensitivity (Temple's criteria) and the severe cases determination (Hy's law criteria) of potential liver toxicity.

There was a bigger numerical impact of the definition is for the Temple's criteria than Hy's law because very few patients had a bilirubin >34 μmol/L. In the general population there was no case with bilirubin greater than 34 μmol/L and ALT > 3 × 26 IU/L and therefore no increase could be observed with ALT > 3 × 66 IU/L. In the tertiary center population the prevalence of Temple criteria changes from 3.87 to 13.56 that is a multiplication by 3.50 and the prevalence of Hy's law from 1.35 to 3.23, that is a multiplication by 2.39. If these figures would be the same for DILI they are clinically very significant for Hy's law cases due to their poor prognostic, even if the impact is slightly lower than for Temple's criteria.

The variability of ULN definition is also a major limitation for the "R ratio" (ALT/ULN)/(Alkaline Phosphatase/ULN) recommended for categorizing liver injury as either hepatocellular or cholestatic (Fontana et al., 2010). It was predictable that a lower ULN will be associated with higher liver disease estimates, but the magnitude of this impact has not been recognized.

4.2. Reference population

We followed most of the Dutta's et al. recommendations for assessing the reference values (Dutta et al., 2009). In our study, the observed ULN-ALT variability could not be due to analytical variability as all the assays were performed by a centralized laboratory using the same analyzers and kits, following the manufacturers' recommendations. We used a reference population of 40 years or older, apparently healthy individuals, with normal body weight without underlying acute or chronic illnesses, with no significant alcohol consumption and no intake of medicines. As previously described we found that the main factors associated with ULN-ALT variability were gender and BMI (Piton et al., 1998; Dutta et al., 2009). We found no significant association with alcohol consumption and medicines prescriptions. For alcohol consumption we previously observed in the same volunteers population that carbohydrate deficient transferrin (CDT) was probably more sensitive and specific than self declared consumption to

Table 2
Prevalence of necro-inflammatory estimates according to ULN ALT definition.

	Low risk population			Tertiary center		
	All	Male	Female	All	Male	Female
	7463	4113	3350	6865	3900	2936
Temple's criteria (ALT > 3 ULN)						
ULN ALT = 26 IU/L	105 (1.41;1.15–1.70)[a]	82 (1.99;1.59–2.47)	23 (0.69;0.44–1.03)	931 (13.56;12.76–14.39)	618 (15.85;14.71–17.03)	313 (10.66;9.57–11.83)
ULN ALT = 66 IU/L	6 (0.08;0.03–1.75)[b]	6 (0.15; 0.05–0.32)	0 (0;0–0.10)[b]	266 (3.87;3.43–4.36)[b]	197 (5.05;4.39–5.79)[b]	69 (2.35;1.83–2.96)[b]
Hy's criteria (ALT > 3 ULN and bilirubin > 34 μmol/L						
ULN ALT = 26 IU/L	0 (0; 0–0.05)	0 (0; 0–0.09)	0 (0;0–0.10)	222 (3.23;2.83–3.68)	171 (4.38;3.76–5.08)	51 (1.74;1.30–2.28)
ULN ALT = 66 IU/L	0 (0; 0–0.05)	0 (0; 0–0.09)	0 (0;0–0.10)	93 (1.35;1.09–1.66)[b]	70 (1.79;1.40–2.26)[b]	23 (0.78;0.50–1.17)[c]

[a] n (%, binomial 95%CI)
[b] Significance of comparisons between the two definitions of ULN-ALT. $P < 0.0001$.
[c] Significance of comparisons between the two definitions of ULN-ALT. $P = 0.001$.

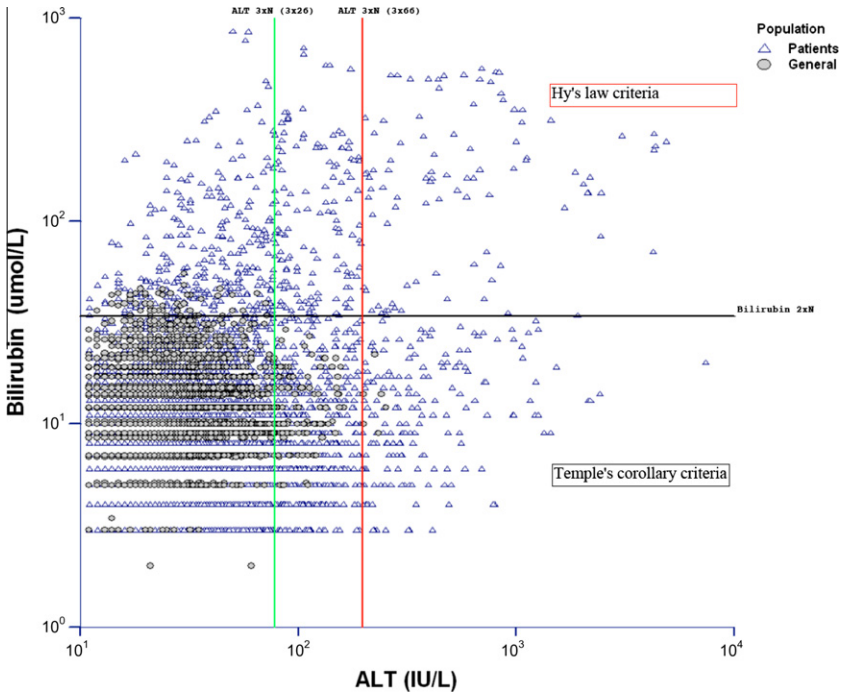

Fig. 1. Impact of the procedures chosen for the ALT upper normal limit (ULN) definition on the prevalence of DILI estimates: Temple's criteria and Hy's law criteria. The green vertical line is the 3-fold ULN-ALT using 26 IU/L as threshold (3 × 26 = 78 IU/L). The red vertical line is the 3-fold ULN-ALT using 66 IU/L as threshold (3 × 66 = 198 IU/L). Patients at the right of the vertical lines fulfilled the "Temple's criteria" for suspicion of liver necrosis or inflammation. Extreme values on both *x* and *y* axis were collapsed in the figure in order to visualize the distribution of the lower values. The horizontal line is the 2-fold bilirubin ULN using 17 micromol/L as threshold (2 × 17 = 34 micromol/L). Patients over this horizontal line and at the right of the vertical lines fulfilled the "Hy's law criteria" for severe necrosis. Triangle are values from the patients hospitalized in the tertiary center. Circle are the values from the subjects representing the low risk population. Using 26 IU/L as ULN, among the patients, 931 (14%) fulfilled "Temple's criteria" and 222 (3%) fulfilled "Hy's criteria". Among the low risk population, 105 subjects (1.4%) fulfilled "Temple's criteria" and none met "Hy's criteria". Using 66 IU/L as ULN, among patients, 266 (4%) had "Temple's criteria" and 93 (3%) had "Hy's criteria". Among low risk population, 6 subjects (0.1%) had "Temple's criteria" and none had "Hy's criteria". (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article.)

identify excessive drinkers (Poynard et al., 2010). In the present study the number of subjects without any risk factors, including normal CDT for eliminating excessive drinkers, was too small for definitive conclusion. A larger study is needed using CDT to test the hypothesis that ULN-ALT will be lower in patients with CDT lower than 1.6%.

It has been suggested for DILI to explore outcome-based reference intervals for ALT, rather than population-based reference values as currently practiced (Dutta et al., 2009). In a general population study a positive association between the ALT concentration, even within normal range (35–40 IU/l), and mortality from liver disease has been observed (Kim et al., 2004). However only validated biomarkers of fibrosis, such as FibroTest, have validated reference intervals which are predictive of morbidity and mortality in chronic liver disease in patients with chronic hepatitis C (Ngo et al., 2006), hepatitis B (Ngo et al., 2008) and alcoholic liver disease (Naveau et al., 2009). ALT levels were not independently correlated with morbidity and mortality in these chronic liver diseases and it is not sure if ALT could be validated by morbidity/mortality outcome in non severe DILI. In addition, in the very rare severe DILI cases it seems also very difficult to validate ALT using liver outcome as so far only bilirubin and prothrombin time have been clearly associated with mortality or transplantation.

4.3. Consensus on reference population ?

From our volunteers of normal weight, no medication use and low alcohol consumption the ULN-ALT was 48 IU/L for male and 38 IU/L for female, as defined by the 95% percentile. These ULN-ALT are similar to those usually recommended, ranging from 30–50 IU/L (Dutta et al., 2009; Weil et al., 2008; Kim et al., 2004; D, 2002). Prati et al. recommended in male < 40 IU/L, in female

Table 3
ULN-ALT in the reference "low risk of ALT increase" low risk population according to recommended criteria[a].

Characteristics of reference population	Number	95% percentile, IU/L (95% CI)
All patients without liver history	7463	55 (53–56)
Male	4113	61 (59–63)
Female	3350	43 (41–44)
Normal body weight (BMI ⩽25 kg/m2)	3586	44 (43–46)
No medication use	4113	56 (53–58)
Low alcohol consumption[b]	5777	54 (53–56)
BMI ⩽25 kg/m^2, no medication use, low alcohol	866	45 (46–53)
Male	528	48 (44–57)
Female	338	38 (34–47)
Carbohydrate deficient transferrin (CDT)<1.6%	749	51 (47–55)
Normal BMI, no medication use, CDT < 1.6%	73	46 (38–92)

[a] All patients had 40 years of age or older; criteria were those of Dutta et al. (2009).
[b] Declared consumption not higher than 10 g/day for female and 20 g/day for male.

< 30 IU/L) using healthy subjects with BMI < 25 kg/m2, serum triglycerides concentrations ⩽200 mg/dL, serum glucose levels ⩽105 mg/dL, and serum total cholesterol concentrations ⩽220 mg/dL (D, 2002). Lee et al. identify recommended lower ULN of ALT in Asian (male < 30, female < 19 IU/L,) using the criteria of Prati et al., modified by the BMI cutoff points for Asian populations (<23 kg/m2), in 665 subjects. This study is very accurate as these volunteers were living liver donors with normal biopsy (Lee et al., 2010). However even this design is not perfect due to the selected population and due to the sampling error of biopsy particularly for the diagnosis of non-alcoholic steato-hepatitis (Ratziu et al., 2005).

Other limitation to our study includes the fact that we used 2 large populations but without identifying the DILI cases among patients with minimal or severe estimates. The ability to identify cases of DILI would improve the risk assessment but would not change the principle of limiting artificial variability between studies. The prevalence of ALT estimates in our low risk population was similar to that of a placebo group (http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/ Guidances/UCM174090.pdf, 0000). Our two populations were useful to demonstrate the impact of an absence of consensus in the definition of ULN for assessing the level of DILI estimates. Starting prospective studies in DILI using ALT ULN without consensus seems hazardous. We also acknowledge that there are countries and hospitals that measure ALT in μkat/L, which should transform results ALT expression in IU/L (to convert the alanine aminotransferase thresholds to μkat/L, multiply by 16.667).

As a single analytical ALT enzymatic method is not used by all laboratories the reference population used to determine the ULN should be described in any scientific publications on DILI prevalence. From the evidence based data an agreed consensus on a reference population seems possible (Pratiet al., 2002; Lee et al., 2010; Ozer et al., 2010). A consensus already exists for separating male and female reference values, excluding patients with overweight, medication uses, and excessive declared alcohol consumption. In the context of DILI a consensus also exists on the predictive value of the change in ALT rather than the absolute level itself.

Three main questions are still debated: 1-which threshold for body mass index defining overweight?, 2-are exclusion criteria using metabolic factors (glucose, triglycerides, cholesterol) too sensitive?, and 3-how to define excessive alcohol consumption, declared or using carbohydrate-deficient transferrin?

4.4. Which recommendations could be made to address arbitrary use of ULN thresholds between research groups?

Two recommendations are simple. ALT assay must be detailed and ALT quantitative results must be given in DILI cases and in the reference population, not just the report of thresholds (Ozer et al., 2010).

A major unsolved methodological issue is the definition of DILI for the assessment of the quantitative performances (such as area under the receiver characteristics curve [AUROC]) of "DILI biomarkers". Ideally the reference criteria for calculating AUROCs would be a liver biopsy both in controls and suspected DILI cases. As DILI is usually defined as an increase of ALT, the true performance of ALT is unknown.

From the evidence based we just know the distribution of ALT among control population at low risk of DILI, with (Lee et al., 2010) or without (Poynard et al., 2010; Prati et al., 2002) biopsy, which is only an imperfect estimate of specificity. The specificity of ALT for DILI in patients without another cause of acute liver disease, varied from zero in patients taking statins for a 100 IU/L threshold (Athyros et al., 2010) to almost 100% in patients taking acetaminophen (more than 10 g/day) for a 500 IU/L threshold. No large studies have been performed to evaluate sensitivity of transaminases.

5. Conclusion

The present study illustrates the limitations of expressing DILIs with reference to ULNs for ALT. The ULN is dependent on the reference population used.

Artificial statistical modifications of the procedures chosen for the upper normal limit definition can change dramatically the prevalence of DILI estimates. A consensus in estimates definitions including the definition of reference populations seems mandatory for DILI studies to prevent misleading conclusions.

Conflict of Interest Statement

Thierry Poynard is the inventor of liver injury biomarkers FibroTest ActiTest with a capital interest in Biopredictive the company marketing the tests. Patents belong to the public organization Assistance Publique Hôpitaux de Paris.

Mona Munteanu and Yen Ngo are full employee of Biopredictive the company marketng FibroTest (FibroSure in USA)

References

Athyros, V.G., Tziomalos, K., Gossios, T.D., Griva, T., Anagnostis, P., Kargiotis, K., Pagourelias, E.D., Theocharidou, E., Karagiannis, A., Mikhailidis, DP.GREACE Study Collaborative Group, 2010. Safety and efficacy of long-term statin treatment for cardiovascular events in patients with coronary heart disease and abnormal liver tests in the greek atorvastatin and coronary heart disease evaluation (GREACE) Study: a post-hoc analysis. Lancet 376, 1916–1922.

Castera, L., Pinzani, M., 2010. Non-invasive assessment of liver fibrosis: are we ready? Lancet 375, 1419–1420.

Dutta, A., Saha, C., Johnson, C.S., Chalasani, N., 2009. Variability in the upper limit of normal for serum alanine aminotransferase levels: a statewide study. Hepatology 50, 1957–1962.

Ferard, G., Piton, A., Messous, D., Imbert-Bismut, F., Frairi, A., Poynard, T., et al., 2006. Intermethod calibration of alanine aminotransferase (ALT) and gamma-glutamyltransferase (GGT) results: application to Fibrotest and Actitest scores. Clin. Chem. Lab Med. 44, 400–406.

Fontana, R.J., Seeff, L.B., Andrade, R.J., Björnsson, E., Day, C.P., Serrano, J., Hoofnagle, J.H., 2010. Standardization of nomenclature and causality assessment in drug-induced liver injury: summary of a clinical research workshop. Hepatology 52, 730–742.

Halfon, P., Imbert-Bismut, F., Messous, D., Antoniotti, G., Benchetrit, D., Cart-Lamy, P., Delaporte, G., Doutheau, D., Klump, T., Sala, M., Thibaud, D., Trepo, E., Thabut, D., Myers, R.P., Poynard, T., 2002. A prospective assessment of the inter-laboratory variability of biochemical markers of fibrosis (FibroTest) and activity (ActiTest) in patients with chronic liver disease. Comp. Hepatol. 1, 3.

Imbert-Bismut, F., Messous, D., Thibaut, V., Myers, R.B., Piton, A., Thabut, D., et al., 2004. Intralaboratory analytical variability of biochemical markers of fibrosis (Fibrotest) and activity (Actitest) and reference ranges in healthy blood donors. Clin. Chem. Lab Med. 42, 323–333.

Kim, H.C., Nam, C.M., Jee, S.H., Han, K.H., Oh, D.K., Suh, I., 2004. Normal serum aminotransferase concentration and risk of mortality from liver diseases: prospective cohort study. BMJ 328, 983.

Lee, J.K., Shim, J.H., Lee, H.C., Lee, S.H., Kim, K.M., Lim, Y.S., Chung, Y.H., Lee, Y.S., Suh, D.J., 2010. Estimation of the healthy upper limits for serum alanine aminotransferase in Asian populations with normal liver histology. Hepatology 51, 1577–1583.

Myara, A., Guéchot, J., Imbert-Bismut, F., Lasnier, E., Piton, A., Voitot, H., Férard, G., 2004. Harmonization of liver enzyme results: calibration for aminotransferases and gamma glutamyltransferase. J. Hepatol. 41, 501–502.

Naveau, S., Gaudé, G., Asnacios, A., Agostini, H., Abella, A., Barri-Ova, N., Dauvois, B., Prévot, S., Ngo, Y., Munteanu, M., Balian, A., Njiké-Nakseu, M., Perlemuter, G., Poynard, T., 2009. Diagnostic and prognostic values of non-invasive biomarkers of fibrosis in patients with alcoholic liver disease. Hepatology 49, 97–105.

NCSS, Hintze.J.L., 2007. User Guide. Number cruncher statistical systems software NCSS, Kaysville, Utah 2007..

Ngo, Y., Munteanu, M., Messous, D., Charlotte, F., Imbert-Bismut, F., Thabut, D., Lebray, P., Thibault, V., Benhamou, Y., Moussalli, J., Ratziu, V., Poynard, T., 2006. A prospective analysis of the prognostic value of biomarkers (FibroTest) in patients with chronic hepatitis C. Clin. Chem. 52, 1887–1896.

Ngo, Y., Benhamou, Y., Thibault, V., Ingiliz, P., Munteanu, M., Lebray, P., Thabut, D., Morra, R., Messous, D., Charlotte, F., Imbert-Bismut, F., Bonnefont-Rousselot, D., Moussalli, J., Ratziu, V., Poynard, T., 2008. An accurate definition of the status of inactive hepatitis B virus carrier by a combination of biomarkers (Fibrotest-Actitest) and viral load. PlosOne 3, e2573.

Ozer, J.S., Chetty, R., Kenna, G., Koppiker, N., Karamjeet, P., Li, D., Palandra, J., Lanevschi, A., Souberbielle, B.E., Ramaiah, S., 2010. Recommendations to qualify biomarker candidates of drug-induced liver injury. Biomark Med. 4, 475–483.

Ozer, J.S., Chetty, R., Kenna, G., Palandra, J., Zhang, Y., Lanevschi, A., Koppiker, N., Souberbielle, B.E., Ramaiah, S.K., 2010. Enhancing the utility of alanine aminotransferase as a reference standard biomarker for drug-induced liver injury. Regul. Toxicol. Pharmacol. 56, 237–246.

Piton, A., Poynard, T., Imbert-Bismut, F., Khalil, L., Delattre, J., Pelissier, E., Sansonetti, N., Opolon, P., 1998. Factors associated with serum alanine transaminase activity in healthy subjects: consequences for the definition of normal values, for selection of blood donors, and for patients with chronic hepatitis C. MULTIVIRC Group. Hepatology 27, 1213–1219.

Poynard, T., 2011. First-line assessment of patients with chronic liver disease with non-invasive techniques and without recourse to liver biopsy. J. Hepatol. 54, 586–587.

Poynard, T., Ratziu, V., Naveau, S., Thabut, D., Charlotte, F., Messous, D., Capron, D., Abella, A., Massard, J., Ngo, Y., Munteanu, M., Mercadier, A., Manns, M., Albrecht, J., 2005. The diagnostic value of biomarkers (SteatoTest) for the prediction of liver steatosis. Comp. Hepatol. 4, 10.

Poynard, T., Munteanu, M., Ngo, Y., Castera, L., Halfon, P., Ratziu, V., Imbert-Bismut, F., Thabut, D., Bourliere, M., Cacoub, P., Messous, D., de Ledinghen, V., 2010. ActiTest accuracy for the assessment of histological activity grades in patients with chronic hepatitis C, an overview using Obuchowski measure. Gastroenterol. Clin. Biol. 34, 388–396.

Poynard, T., Lebray, P., Ingiliz, P., Varaut, A., Varsat, B., Ngo, Y., Norha, P., Munteanu, M., Drane, F., Messous, D., Bismut, F.I., Carrau, J.P., Massard, J., Ratziu, V., Giordanella, J.P., 2010. Prevalence of liver fibrosis and risk factors in a general population using non-invasive biomarkers (FibroTest). BMC Gastroenterol. 10, 40.

Prati, D., Taioli, E., Zanella, A., Della Torre, E., Butelli, S., Del Vecchio, E., Vianello, L., Zanuso, F., Mozzi, F., Milani, S., Conte, D., Colombo, M., Sirchia, G., 2002. Updated definitions of healthy ranges for serum alanine aminotransferase levels. Ann. Intern. Med. 2 (137), 1–10.

Ratziu, V., Charlotte, F., Heurtier, A., Gombert, S., Giral, P., Bruckert, E., Grimaldi, A., Capron, F., 2005. LIDO Study Group. Sampling variability of liver biopsy in nonalcoholic fatty liver disease. Gastroenterology 128, 1898–1906.

Watkins, P.B., Seligman, P.J., Pears, J.S., Avigan, M.I., Senior, J.R., 2008. Using controlled clinical trials to learn more about acute drug-induced liver injury. Hepatology 48, 1680–1689.

Weil, J.G., Bains, C., Linke, A., Clark, D.W., Stirnadel, H.A., Hunt, C.M., 2008. Background incidence of liver chemistry abnormalities in a clinical trial population without underlying liver disease. Regul. Toxicol. Pharmacol. 52, 85–88.

<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/ Guidances/UCM174090.pdf>.