Alanine Aminotransferase Results Differ by Analyzer Manufacturer in a National Integrated Health Setting, 2012–2017

Lauren A. Beste, MD, MSc; Michael Icardi, MD; Christine M. Hunt, MD, MPH; Ina Gylys-Colwell, MS; Elliott Lowy, PhD; Leslie Taylor, PhD; Timothy R. Morgan, MD; Michael F. Chang, MD; Marissa M. Maier, MD; Ramsey Cheung, MD

- **Context.**—Disease guidelines specify universal alanine aminotransferase (ALT) thresholds for clinical decision-making, yet the effect of variability among ALT analyzers remains unclear.

- **Objective.**—To compare ALT results from different analyzers from 2012–2017.

- **Design.**—Veterans Health Administration (VHA) laboratories perform external ALT proficiency testing using standardized College of American Pathologists (CAP) samples in analyzers by 5 manufacturers. In this operational analysis, we evaluated 22 950 ALT values from 80 independent CAP samples tested at 223 laboratories. Using mixed effects modeling, we estimated the association between analyzer manufacturer and CAP outcome, adjusting for manufacturer, facility, and calendar year. We performed subgroup analyses on CAP samples with overall means near clinical guideline-specified thresholds, including less than 50 U/L (n = 10) and less than 35 U/L (n = 5).

- **Results.**—The VHA used Abbott Laboratories (n = 3175; 14%), Beckman Coulter Diagnostics (n = 8723; 38%), Roche Diagnostics (n = 2595; 11%), Siemens Healthineers USA (n = 5713; 25%), and Vitros/Ortho Clinical Diagnostics (n = 2744; 12%) analyzers. The CAP samples (n = 80 samples, n = 22 950 tests) covered a wide range of mean ALT values (21–268 U/L). The average difference in mean ALT value per sample between the highest-reading and lowest-reading manufacturers was 15.4 U/L (SD = 1.8) for the 10 samples with mean ALT less than 50 U/L, and it was 10.4 U/L (SD = 3.6) overall (n = 80). In linear mixed effects modeling, we found statistically significant differences in ALT values between the different manufacturers in each year.

- **Conclusions.**—We found statistically and clinically meaningful differences between analyzers across the ALT spectrum in each year, including at ALT levels lower than 50 U/L and lower than 35 U/L. Universal ALT thresholds should be avoided as a trigger for clinical action until differences between analyzers can be resolved.

(Arch Pathol Lab Med. 2020;144:748–754; doi: 10.5858/arpa.2018-0622-OA)

Alanine aminotransferase (ALT) is a marker of hepatocyte injury widely used to screen for and monitor liver disease.\(^1,2\) Alanine aminotransferase features prominently in clinical algorithms for detecting and managing liver disease and for triggering subspecialty referral. For example, US and European guidelines recommend evaluation for suspected nonalcoholic fatty liver disease among individuals with elevated liver enzymes and fatty liver,\(^3,4\) which is the most common cause of incidental liver enzyme elevations in patients presenting to primary care.\(^5,6\) The US guidelines recommend using ALT elevations to determine the need for antiviral therapy in patients with chronic hepatitis B virus infection.\(^7\) Expressed as a multiple (or “fold”) of upper limit of normal (ULN), ALT elevations are commonly used to identify liver injury in clinical trials as well as in drug-induced liver injury.
Known variability exists in laboratory determination of ALT. The American College of Gastroenterology (ACG) specifies a range of “normal” ALT values based on published results of healthy individuals in blood donor populations, healthy liver transplant donors, and the National Health and Nutrition Examination Survey. Recent ACG guidelines identify the ULN for ALT as 32 to 33 U/L (males) and 20 to 25 U/L (females) (to convert to microkatal per liter, multiply by 0.0167), irrespective of ALT assay or analyzer. Similarly, the American Association for the Study of Liver Diseases recommends that chronic hepatitis B patients with ALT exceeding 2 times ULN (defined for the Study of Liver Diseases recommends that chronic hepatitis B virus antiviral treatment. As 35 U/L [males] and 25 U/L [females]) be considered for hepatitis B patients with ALT exceeding 2 times ULN (defined for the Study of Liver Diseases recommends that chronic hepatitis B virus antiviral treatment.

Variability in ALT measurement by analyzer manufacturer and over time has important implications if universal “normal” cutoffs are to be applied. Alanine aminotransferase variability across analyzer platforms in the United States has received only limited study. Concerningly, a recent large national Canadian laboratory study reported a 25% coefficient of variation in ALT results across laboratories (n = 40) and a 22% variation in the ALT reference range across analyzer manufacturers using a single standardized ALT sample.

As the largest integrated US health care network, the Veterans Health Administration (VHA) presents a unique opportunity to examine the performance of ALT analyzers on a national scale. The VHA has more than 9 million enrollees, 170 medical centers, and 1061 outpatient sites of care. All VHA laboratories operate independently and are Clinical Laboratory Improvement Amendments (CLIA) certified. All VHA laboratories perform externally validated laboratory proficiency testing through the College of American Pathologists (CAP) 3 times per year using standardized ALT samples. Using CAP test results from 2012-2017, we sought to: (1) compare ALT results produced by 5 different analyzer manufacturers; and (2) examine whether trends in analyzer performance remained stable over time.

MATERIALS AND METHODS

Data Source

The VHA Office of Pathology and Laboratory Medicine oversees all VHA laboratories and requires them to meet the rigorous accreditation standards of the CAP, in addition to the US federal standards of the CLIA for acceptable analytic performance. As part of its accreditation procedures, the CAP independently prepares blinded aliquots of 5 ALT samples and supplies them to each participating VHA and non-VHA laboratory 3 times per year, for a total of 15 CAP samples per year. CAP testing was performed on every ALT analyzer operating within the VHA at the time of testing. Testing was carried out in compliance with federal regulations and adhered to the analyzer manufacturers’ instructions. The VHA Office of Pathology and Laboratory Medicine provided data on ALT values and analyzer manufacturers for CAP testing performed from 2012-2017 within the VHA system. It also provided characteristics of the laboratories for 2017, including VHA-defined complexity of the parent facility (a surrogate for tertiary status), and the number of full-time equivalent clinical pathologists on staff. Volume of ALT tests performed per facility varied throughout 2012-2017, an overall range of 40 to 15 000 test results per facility per year. We included an interaction term for year and manufacturer to account for the possibility that the analyzer model or other technical factors, such as reagents, might have changed over time within manufacturers. We report average adjusted mean ALT values for each manufacturer across all years and their associated significance tests. We tested the significance of the fixed effects and interaction term using F-tests. To determine whether differences between analyzers were present at all levels of ALT, we repeated the linear mixed effects modeling for subgroups deemed clinically meaningful (in the determination of the normal range), including less than 50 U/L and over 35 U/L. We used the VHA-wide mean ALT for each CAP sample to select these 2 subgroups because the “true” ALT values were unknown.

We also identified the CAP sample with the lowest overall ALT mean. This sample was selected because it was closest to the range for ULN recommended in ACG practice guidelines for healthy males (29 to 33 U/L). The proportion of VHA results exceeding the recommended ULN threshold of 33 U/L was calculated overall and by analyzer manufacturer for this single CAP sample.

Analyses were performed using SAS version 9.4 (Research Triangle Park, North Carolina). All means and descriptive statistics are unadjusted, unless specified. All years are calendar years, unless specifically indicated as fiscal years.

RESULTS

VHA Facilities and Analyzer Manufacturers

Although the number of laboratories and analyzer machines in use varied throughout 2012–2017, an overall...
223 laboratories in 140 parent facilities completed CAP testing of 22,950 specimens. The ALT measurements were performed on analyzers using standard procedures from 5 manufacturers: Abbott Laboratories (n = 3,175 [14%]), Beckman Coulter Diagnostics (n = 8,723 [38%]), Roche Diagnostics (n = 2,595 [11%]), Siemens Healthineers USA (n = 5,713 [25%]), and Vitros/Ortho Clinical Diagnostics (n = 2,744 [12%]). In total, we excluded 345 CAP results (1.5%) that were classified as nonviable or were rejected because of technical issues (Table 1).

In 2017, results from 5 CAP samples were processed by 369 machines in 210 laboratories (Table 2; data provided by VHA National Pathology and Laboratory Medicine Services). For inpatient and outpatient care, each VHA parent facility performed a mean of 61,481 (SD = 39,457) ALT tests the preceding year; 40 of 141 sites (28%) were classified as tertiary (highest complexity). The locations of the 5 types of analyzers varied in the 9 VHA regions in 2017, ranging from 4 regions with the Vitros/Ortho analyzer up to 8 regions for Beckman (Table 2).

### Differences in CAP Results by Analyzer Manufacturer

From the 80 individual CAP samples, the mean VHA-wide ALT results ranged from 28.5 to 236.3 U/L per sample (Figure 1; Supplemental Table [see supplemental digital content at www.archivesofpathology.org in the June 2020 table of contents]). These ALT means were nearly identical to the US published means for these same CAP samples (Supplemental Table).

Across all 80 CAP samples, we found a mean difference of 10.4 U/L (SD = 3.6) between the highest- and lowest-reading manufacturers. Vitros/Ortho analyzers consistently produced higher results than analyzers by other manufacturers (Figure 2; Table 3).

### Table 1. College of American Pathologists (CAP) Test Outcomes on Samples Tested 2012–2017, by Alanine Aminotransferase (ALT) Analyzer Manufacturer (n = 80 Samples)a

|        | Abbott | Beckman | Roche | Siemens | Vitros/Ortho | Total |
|--------|--------|---------|-------|---------|--------------|-------|
| Viable CAP results, No. (%) | 3,175 (14) | 8,723 (38) | 2,595 (11) | 5,713 (25) | 2,744 (12) | 22,950 |
| Uninterpretable/missing, No. (%) | 25 (0.8) | 92 (1.0) | 65 (2.5) | 122 (2.1) | 21 (0.8) | 325 |
| Laboratories, No. (%) | 42 (17) | 89 (35) | 26 (10) | 66 (26) | 29 (12) | 252b |
| Analyzer machines, No. (%) | 67 (16) | 144 (34) | 49 (11) | 117 (27) | 52 (12) | 429 |

a Abbott Laboratories; Beckman Coulter Diagnostics; Roche Diagnostics; Siemens Healthineers USA; Vitros/Ortho Clinical Diagnostics.

b A total of 29 of 223 laboratories added or changed manufacturers during a 6-year period.

### Table 2. Facility-Level Characteristics of Veterans Health Administration Laboratories in 2017, by Alanine Aminotransferase (ALT) Analyzer Manufacturera

|        | Abbott | Beckman | Roche | Siemens | Vitros/Ortho | Total |
|--------|--------|---------|-------|---------|--------------|-------|
| parent facility, No. | 26 | 44 | 17 | 41 | 13 | 141 |
| ALT tests performed per facility, mean (SD) | 58,663 (36,512) | 72,804 (46,326) | 63,438 (49,241) | 49,366 (28,082) | 64,444 (28,821) | 61,481 (39,457) |
| Pathologist clinical FTE, mean (SD) | 2.5 (2.3) | 2.5 (1.6) | 2.6 (1.6) | 2.2 (1.5) | 2.5 (1.2) | 2.4 (1.7) |
| Parent facility complexity, n (%) | | | | | | |
| High complexity | 6 (23) | 13 (30) | 9 (53) | 8 (20) | 4 (31) | 40 (28) |
| High complexity | 3 (12) | 9 (20) | 0 (0) | 5 (12) | 4 (31) | 21 (15) |
| High complexity | 6 (23) | 9 (20) | 2 (12) | 13 (32) | 1 (8) | 31 (22) |
| Moderate complexity | 5 (19) | 6 (14) | 1 (6) | 6 (15) | 2 (15) | 20 (14) |
| Waived | 6 (23) | 7 (16) | 5 (29) | 8 (20) | 2 (15) | 28 (20) |
| Missing | 0 (0) | 0 (0) | 0 (0) | 1 (2) | 0 (0) | 1 (1) |
| Geographic region, n (%) | | | | | | |
| New England | 8 (31) | 0 (0) | 0 (0) | 0 (0) | 0 (0) | 8 (6) |
| Mid-Atlantic | 0 (0) | 10 (23) | 0 (0) | 8 (20) | 0 (0) | 18 (13) |
| Midwest | 2 (8) | 1 (2) | 1 (6) | 18 (44) | 0 (0) | 22 (16) |
| Plains | 6 (23) | 1 (2) | 2 (12) | 0 (0) | 4 (31) | 13 (9) |
| South Atlantic | 5 (19) | 5 (11) | 6 (35) | 8 (20) | 0 (0) | 24 (17) |
| South | 1 (4) | 2 (5) | 1 (6) | 1 (2) | 5 (38) | 10 (7) |
| Central Gulf | 2 (8) | 12 (27) | 0 (0) | 0 (0) | 1 (8) | 15 (11) |
| Mountain West | 2 (8) | 3 (7) | 0 (0) | 5 (12) | 3 (23) | 13 (9) |
| West Coast/Philippines | 0 (0) | 9 (20) | 7 (41) | 1 (2) | 0 (0) | 17 (12) |

Abbreviation: FTE, full-time equivalent.

a Analyzers: Abbott Laboratories; Beckman Coulter Diagnostics; Roche Diagnostics; Siemens Healthineers USA; Vitros/Ortho Clinical Diagnostics.

b A total of 139 Veterans Health Administration parent facilities in 2017; 2 facilities used 2 different manufacturers.

c New England: CT, MA, ME, NH, RI, VT. Mid-Atlantic: NJ, NY, PA. Midwest: IL, IN, MI, OH, WI. Plains: IA, KS, MN, MO, ND, NE, SD. South Atlantic: DC, DE, FL, GA, MD, NC, PR, SC, VA, WV. South: AL, KY, MS, TN. Central Gulf: AR, LA, OK, TX. Mountain West: AZ, CO, ID, MT, NM, NV, UT, WY. West Coast/Philippines: AK, CA, HI, OR, WA, Philippines.
Figure 1. Range in mean alanine aminotransferase (ALT) between highest-reading and lowest-reading analyzer manufacturers on standardized College of American Pathologists (CAP) testing, 2012–2017 (n = 80 samples).

Figure 2. Mean alanine aminotransferase (ALT) on standardized College of American Pathologists (CAP) testing, by manufacturer (n = 80 samples). Abbott Laboratories; Beckman Coulter Diagnostics; Roche Diagnostics; Siemens Healthineers USA; Vitros/Ortho Clinical Diagnostics.
different analyzer manufacturers (Supplemental Table). Within this subgroup, we found a mean difference of 15.4 U/L (SD = 1.8) between the highest- and lowest-reading manufacturers. Across different analyzer manufacturers, mean ALT for the 5 individual CAP samples in the less than 35 U/L subgroup ranged from 28.5 to 34.2 U/L. In this subgroup, we found a mean difference of 16.4 U/L (SD = 1.2) between the highest- and lowest-reading manufacturers. In both the less than 35 U/L and less than 50 U/L ranges, Vitros/Ortho machines produced higher results than other analyzers, and Roche produced the lowest (Figures 2 and 3; Table 3; Supplemental Table).

In adjusted models, the interaction between year and manufacturer was statistically significant ($P < .001$), indicating that differences in ALT values between manufacturers differed by year. Within each given year, there were statistically significant differences in ALT values between the different manufacturers ($P < .001$ for all years). Unadjusted mean results from Roche analyzers were consistently lower and Vitros/Ortho consistently higher compared with the other analyzers.

These statistically significant differences between manufacturers persisted after restricting our sample population to:

- (1) samples with mean ALT less than 50 U/L (all $P < .001$), and
- (2) samples with mean ALT less than 35 U/L (all $P < .001$), and after excluding results from Vitros/Ortho.

### Comparison of VHA ALT Results to ACG Practice Guideline ALT ULN Thresholds for Healthy Males

The CAP sample with the lowest mean ALT (28.5 U/L [SD = 5.9]) was selected for comparison with ALT elevations exceeding the normal thresholds recommended in published guidelines. Across all 5 analyzers, 43% of 227 (19%) of ALT results exceeded the ACG threshold of ULN (33 U/L) for this CAP sample and would be considered an elevated ALT in males. Of the 227 analyzers testing this CAP specimen, Abbott, Beckman, and Roche analyzers did not produce any ALT results higher than the 33 U/L threshold; 11 of 47 (23%) ALT values were above 33 U/L for Siemens analyzers, and 31 of 31 (100%) ALT values were above this threshold for Vitros/Ortho analyzers.

### DISCUSSION

Alanine aminotransferase analyzers produced by different manufacturers yield statistically different results that can be clinically relevant, particularly at lower ALT ranges. After controlling for the effects of facility and year in a national health care system, we found a statistically significant association between manufacturer and ALT proficiency testing results during our 6-year analysis. Among samples with a mean ALT lower than 50 U/L, we noted an average 15.4 U/L difference between the highest- and lowest-reading analyzers. We additionally applied the ALT thresholds recommended in US liver disease guidelines to our data and found substantial differences by analyzer manufacturer. For example, the CAP sample with the lowest national mean result (28.5 U/L) produced a mean ALT of 24.3 U/L on Roche analyzers and ALT 41.1 U/L on Vitros/Ortho analyzers. Applying the ACG “normal” threshold of ALT 29 to 33 U/L for healthy males to proficiency testing data, we found that nearly 1 in 5 results from this “normal” ALT sample would have been classified as abnormal (>33 U/L) in clinical practice. Although proficiency test samples for ALT are not necessarily commutable with human specimens and therefore cannot be extrapolated literally to real-world results, our findings raise significant concern that cross-platform variability may also occur in clinical practice. At a minimum, our findings suggest a need to further investigate the source of the observed variability and to consider the need for standardization before applying universal ALT cutoffs in clinical practice.

With guidelines recommending the evaluation of unexplained ALT elevations, our findings suggest that many patients are at risk for misclassification as having “abnormal” ALT simply as a result of analyzer characteristics. Misclassification could potentially launch a negative cascade of events, including needless anxiety and uncertainty for patients and families, costly additional testing, and unnecessary specialist referral and treatment. Clinical pathologists have already highlighted their concerns for the use of a universal ALT threshold (with its potential for overdiagnosis of liver disease) and the importance of pathologist input in guideline development. Our results suggest that analyzer-specific normal ALT ranges might be more appropriate than an absolute cutoff for normal range, at least until cross-platform analyzer differences can be further explored.

---

**Table 3. College of American Pathologists (CAP) Results for All Tests 2012–2017, Samples <50 U/L, Samples <35 U/L, and Lowest Single Sample**

| CAP samples included, No. | Mean All <50 U/L | Mean All <35 U/L | Mean Single Alanine Aminotransferase Sample |
|---------------------------|-----------------|-----------------|--------------------------------------------|
| All Tests, No.            | 22,950          | 2,528           | 1,254                                      |
| Mean, U/L                 | 116.8           | 38.4            | 31.0                                       |
| SD                        | 51.8            | 9.5             | 6.0                                        |
| Abbott                    | 3,175           | 341             | 168                                        |
| Mean, U/L                 | 118.2           | 35.3            | 27.8                                       |
| SD                        | 53.7            | 8.1             | 3.1                                        |
| Beckman                   | 8,723           | 991             | 494                                        |
| Mean, U/L                 | 114.0           | 37.2            | 29.6                                       |
| SD                        | 50.3            | 8.1             | 2.8                                        |
| Roche                     | 2,595           | 296             | 144                                        |
| Mean, U/L                 | 111.7           | 34.4            | 26.7                                       |
| SD                        | 51.0            | 8.0             | 2.4                                        |
| Siemens                   | 5,713           | 573             | 282                                        |
| Mean, U/L                 | 119.9           | 37.9            | 30.3                                       |
| SD                        | 53.0            | 8.9             | 4.4                                        |
| Vitros/Ortho              | 2,744           | 327             | 166                                        |
| Mean, U/L                 | 122.1           | 49.8            | 43.1                                       |
| SD                        | 52.0            | 8.6             | 4.5                                        |

*Key findings appear in boldface font.*

**Abbott Laboratories; Beckman Coulter Diagnostics; Roche Diagnostics; Siemens Healthineers USA; Vitros/Ortho Clinical Diagnostics.**
When examining the broader spectrum of ALT data from 29 to 236 U/L, we found a mean difference of 10.4 U/L (SD = 3.6) between the highest- and lowest-reading manufacturer means for each sample. Our ALT findings mirror a recent Canadian national study (performed with pooled human blood samples) as well as a state-wide study in Indiana; Siemens analyzers were “biased high” compared with Abbott, Beckman, and Roche analyzers, which were “biased low.” We found Vitros/Ortho analyzers to be the highest-reading, followed by Siemens, Abbott, Beckman, and Roche. When we excluded the results of the Vitros/Ortho analyzer, there was greater consistency between the other 4 manufacturers, yet statistically significant variation persisted. These observed differences in ALT values across analyzer manufacturers highlight the need for provider awareness of the ALT reference range when assessing ALT test results. In patients with markedly elevated ALT, between-manufacturer differences of less than 20 U/L are unlikely to influence management but could potentially cause misclassification at lower ALT levels.

Among our study’s strengths are its national scope and use of externally certified, high-quality ALT proficiency tests during 6 years. The VHA proficiency testing results from our integrated national health care system are nearly identical to those from US-wide proficiency data, supporting their reliability. All VHA sites are certified by CAP and adhere to analyzer manufacturers’ instructions. We examined more than 22 000 CAP samples over a wide range of ALT values—far more than any previously published work—and accounted for facility-level effects and year. We accounted for potential changes in analysis technique over time by including an interaction term representing the product of analyzer and year.

Our conclusions should be interpreted within the context of several limitations. Most importantly, although the CAP publishes the national mean and other statistics for its standardized specimens, the objective “target” result for each CAP sample is unknown. This lack of a gold standard precludes an assessment of the relative superiority of one analyzer manufacturer over another. Without an international certified ALT reference standard and ALT assay standardization (eg, use of coenzyme pyridoxal 5’-phosphate), we cannot provide insight into the many other factors that may influence the determination of what constitutes a “normal” ALT, such as patient characteristics (eg, age, sex, ethnic background), and laboratory factors, such as ALT assay reagents. In the future, an international certified reference standard for ALT could be used to investigate and address the cross-platform variability we observed. A second limitation is that our analysis is restricted to the 5 ALT analyzer manufacturers used in the VHA system through 2017. Future work is needed to assess the performance of other manufacturers. Third, any changes in analyzer technology after 2017 are not reflected in our results. Finally, proficiency test data for ALT are not known to be commutable with true human samples for multiple potential reasons, including the use of nonhuman matrix or modification for stability. Study is needed to assess the commutability of reference materials used in external quality assurance for ALT. Future VHA studies in ALT quality control will be explored—for example, to determine whether each manufacturer’s calibration is determined from a known absorptivity value in the indicator reaction.

In conclusion, significant ALT measurement differences occur by analyzer manufacturer, particularly with an ALT in or near the “normal” range. Use of universal thresholds to define the “normal” ALT values will potentially result in some patients being misclassified as having ALT elevations, with the possibility for unnecessary medical intervention due to measurement differences. To address ALT measurement differences by analyzer, manufacturers could standardize their analyzers against an objective reference population, as has been attempted in Europe. In the United States, the Food and Drug Administration has recently started the Food and Drug Administration–Systemic Harmonization and Interoperability for Enhancement of Laboratory Data (SHIELD) program, a pilot for inclusion of the unique device identifier for each lab result as

![Figure 3. Mean alanine aminotransferase (ALT) on standardized College of American Pathologists (CAP) testing among samples with US mean less than 50 U/L, by manufacturer (n = 10 samples). Abbott Laboratories; Beckman Coulter Diagnostics; Roche Diagnostics; Siemens Healthineers USA; Vitros/Ortho Clinical Diagnostics.](image-url)
This would allow health care organizations to implement decision support and reference ranges based on the type of device used. Finally, over and above laboratory harmonization and standardization, continued work is required to determine what constitutes “true normal” ALT in healthy individuals.

References
1. Dufour DR, Lott JA, Nolte FS, Gretch DR, Koff RS, Seeff LB. Diagnosis and monitoring of hepatic injury: I: performance characteristics of laboratory tests. *Clin Chem* 2000;46(12):2027–2049.
2. Kwo PY, Cohen SM, Lim JK. ACG clinical guideline: evaluation of abnormal liver chemistries. *Am J Gastroenterol* 2017;112(1):18–35.
3. Chalasani N, Younossi Z, Lavine JE, et al. The diagnosis and management of nonalcoholic fatty liver disease: practice guidance from the American Association for the Study of Liver Diseases. *Hepatology* 2018;67(1):328–357.
4. European Association for the Study of the Liver (EASL); European Association for the Study of Diabtes (EASD); European Association for the Study of Obesity (EASO). EASL-EASD-EASO Clinical Practice Guidelines for the management of non-alcoholic fatty liver disease. *J Hepatol*. 2016;64(6):1388–1402.
5. Armstrong MJ, Houlihan DD, Bentham L, et al. Presence and severity of non-alcoholic fatty liver disease in a large prospective primary care cohort. *J Hepatol*. 2012;56(1):234–240.
6. Rinella ME. Nonalcoholic fatty liver disease: a systematic review. *JAMA*. 2015;313(22):2263–2273.
7. Terrault NA, Lok ASF, McMahon BJ, et al. Update on prevention, diagnosis, and treatment of chronic hepatitis B: AASLD 2018 hepatitis B guidance. *Hepatology*. 2018;67(4):1556–1599.
8. Dutta A, Saha C, Johnson CS, Chalasani N. Variability in the upper limit of normal for serum alanine aminotransferase levels: a statewide study. *Hepatology* 2009;50(6):1957–1962.
9. Adeli K, Higgins V, Seccombe D, et al. National survey of adult and pediatric reference intervals in clinical laboratories across Canada: a report of the CSCC Working Group on Reference Interval Harmonization. *Clin Biochem*. 2017;50(16-17):925–935.
10. US Department of Veterans Affairs. Veterans Health Administration: providing health care for veterans. https://www.va.gov/health/. Accessed July 5, 2018.
11. Medicare, Medicaid and CLIA programs; regulations implementing the clinical laboratory improvement. *Fed Regist*. 1992;57(40):7002–7186.
12. College of American Pathologists. Standards for laboratory accreditation. In: Laboratory Quality Solutions. 2017. https://ecls.cap.org/ecls/ShowProperty?nodePath=UCMCON/Contribution%20Folders/WebApplications/shared-assets/lap-standards.pdf. Accessed March 30, 2019.
13. Miller WG, Jones GRD, Horowitz GL, Weykamp C. Proficiency testing/external quality assessment: current challenges and future directions. *Clin Chem*. 2011;57:1670–1680.
14. Panteghini M, Adeli K, Ceriotti F, Sandberg S, Horvath AR. American liver guidelines and cutoffs for “normal” ALT: a potential for overdiagnosis. *Clin Chem*. 2017;63(7):1196–1198.
15. Goossens K, Van Uytpleeghe K, Thienpont LM. Trueness and comparability assessment of widely used assays for 5 common enzymes and 3 electrolytes. *Clin Chim Acta*. 2015;442:44–45.
16. Ceriotti F, Schumann G, Panteghini M. Redefining reference limits needs more attention to the analytical aspects. *Liver Int*. 2006;26(9):1155–1156.
17. Vesper HW, Miller WG, Myers GL. Reference materials and commutability. *Clin Biochem Rev*. 2011;32(6):139–147.
18. Kilpatrick ES, Sandberg S. An overview of EFLM harmonization activities in Europe. *Clin Chem Lab Med*. 2018;56(10):1591–1597.
19. Jansen RTP, Cobbaert CM, Weykamp C, Thelen M. The quest for equivalence of test results: the pilgrimage of the Dutch Calibration 2.000 program for metrological traceability. *Clin Chem Lab Med*. 2018;56(10):1673–1684.
20. National Evaluation System for Health Technology Coordinating Center (NESTcc). Developing and implementing sustainable real-world evidence (RWE) infrastructure for in vitro diagnostics (IVDs) through Systemic Harmonization and Interoperability for Enhancement of Laboratory Data (SHIELD). https://nestcc.org/demonstration-projects/. Accessed November 1, 2018.