Effect of Storage Temperature and Time on Stability of Liver Enzymes in Blood Serum

Shadi Kolahdoozan, MD, MPH; Sadaf G. Sepanlou, MD, PhD; Maryam Sharafkhah, PhD; Elaheh Shaker, MD; Ameneh Shayanrad, MSc; Reza Malekzadeh, MD; Shahin Merat, MD; Hossein Poustchi, MD

1Digestive Diseases Research Center, Digestive Diseases Research Institute, Tehran University of Medical Sciences, Tehran, Iran
2Faculty of medicine, Tehran University of Medical Sciences, Tehran, Iran

Abstract

Background: It is increasingly common to collect and store specimens for future unspecified research. However, the effects of prolonged storage on the stability and quality of analytes in serum have not been well investigated. We aimed to determine whether the stability of liver enzymes extracted from frozen bio-samples stored at the baseline is affected by storage conditions.

Methods: A total of four liver enzymes in the sera of 400 patients were examined following storage. After detert-mining the baseline measurements, the serum of each patient was aliquoted and stored at −70°C for three and six months, as well as one, two, and five years after collecting the original sample. The percent change from baseline measurements was calculated both statistically and clinically. Linear models were also used to correct the results of the samples based on the time they were frozen.

Results: In almost all samples, liver enzymes were detectable until two years after the baseline, while in a significant proportion of samples, enzymes were not ultimately detectable five years after the baseline. Linear regression analysis on log-transformed levels of enzymes shows that the performance is acceptable until one year after the baseline. The performance of the prediction model declines substantially two and five years after the baseline, except for GGT.

Conclusion: Long-term storage of serum samples significantly decreases the concentration of the liver enzymes from the baseline, except for GGT. It is not recommended to store samples for more than two years, as liver enzymes are not detectable afterwards.

Keywords: Enzymes, Liver, Storage, Temperature, Time


Received: November 2, 2019, Accepted: January 26, 2020, ePublished: May 1, 2020.

Introduction

Fluid bio-specimens (e.g., serum, plasma, etc.) are widely used in clinical practice and scientific biomedical research. It is common for them to be collected and stored for use at later times and various locations for research purposes.1

The stability of the stored specimens and the influence of controllable parameters (e.g. temperature, thawing, additives, etc) have been investigated in short-term storage.2,3 However, the effects of prolonged storage on frozen serum analytes are not well defined yet.4,5 It has been suggested that long-term storage of specimens at −70°C can compromise the stability of some of the serum analytes.6-9 Samples might undergo multiple freeze-thawing as the result of non-optimal storage conditions, which could be among the most common causes of deterioration in sample quality.10-11 In addition, even under optimal conditions, the storage time might impact the sample quality.9,12 Consequently, these factors can end in partial loss of analytes in the specimens. Hence, it is pivotal to consider these effects in the design of studies and to choose the most stable analytes in long-term frozen and stored samples.

As predictors of liver13 and non-liver-related morbidity and mortality, especially higher total and cardiovascular mortality,14 we assessed long-term effects of storage time on the stability of hepatobiliary enzymes, including γ-glutamyltransferase (GGT), alanine aminotransferase (ALT), aspartate aminotransferase (AST), and alkaline phosphatase (ALP), stored at -70°C, over a time period of five years. Liver enzymes were measured at baseline as well as three and six months, as well as one, two, and five years after storage at baseline. We used stored serum samples in the setting of the repeated measurements phase (2008-2012) in “Golestan Cohort Study” (GCS) (https://decg2.cancer.gov/gemshare/studies/GCS/).15

Material and Methods

Patients Selection

This project is part of the GCS, a prospective study in northeast of Iran, on 50000 participants aged 40–75 years old, who were recruited between 2004 and 2008. All inhabitants in rural areas of Golestan province and
a representative sample selected through systematic clustering in urban areas were recruited in GCS. From 2010 to 2012, a repeated measurement was performed on 11,428 participants randomly selected from the entire cohort. Blood samples were collected from all participants in this phase. Of 400 participants randomly selected out of the entire participants in the repeated measurement phase, serum levels of liver enzymes were measured in fresh blood samples at baseline. The remaining samples were frozen and stored. The stored samples were thawed at 3 months, 6 months, one year, two years, and subsequently five years after collecting the original sample at baseline and serum liver enzymes were measured.

Collection of Human Serum Sample
Each fasting participant donated a 10.0-mL red top Vacutainer (BD Vacutainer Systems, Becton-Dickinson, and Plymouth, UK) tube of whole blood, containing no anticoagulant or preservative. Each tube was left in an upright position for 30 minutes at room temperature, allowed to complete clot followed by centrifuging at 3500 rpm for 10 minutes to isolate serum in a consistent manner to minimize pre-analytic variability. Serum samples were examined visually for hemolysis and lipemia to prevent probable interference.

Resulting serum was divided into 10 equal aliquots, 1.0 mL each, and collected in sterile plastic tubes per donor. One aliquot was stored at 4°C for baseline measurement and as a reference group. The other aliquots were labeled as 3, 6, 12, 24, and 60 months and kept frozen at -70°C until use for further assessment. The time period between venipuncture, serum separation, and storage of the aliquots was less than 48 hours. Overall, we measured 2800 aliquots out of 4000 stored aliquots using Olympus kits. The frozen samples were separately thawed by placing them at room temperature with slight agitation for each thawing endpoint. All measurements were performed by the Olympus AU 400 auto-analyzer equipment.

To eliminate the effect of different methods of measurement, we chose the same testing equipment and methods and even the same reagents in follow-up measurements as those used at the baseline. However, the effect of reagents of different batches could not be eliminated as it would not be feasible to obtain the same batch each time or to keep reagents for five years, which might lead to more errors. Indeed, all of these parameters are tested routinely in a clinical setting and there are standard operating procedures for them. All of these standards ensured consistency in our measurements.

Statistical Analysis
As mentioned above, we measured the levels and trends of liver enzymes, namely ALT, AST, ALP, and GGT at baseline of the study and repeated it three months, six months, one, two, and five years after the baseline. The assumption of normality of the distribution of liver enzymes was tested using Shapiro-Wilk test and normal probability plot. Due to non-normal distribution of data, all data were presented as median with 25th and 75th percentiles at all time points, and log transformation was done for all liver serum enzymes in all analyses. The assumption of normality was valid for log-transformed variables using normal probability plot. For each time point, all enzyme levels were compared with corresponding levels in fresh blood at baseline using paired sample t-test. According to Bonferroni correction for multiple comparisons, P values < 0.01 (0.05/5 = 0.01) were considered statistically significant. Time trend was tested using generalized least square (GLS) model with random effect. Normal distribution was considered for random components in the model.

The percent change (PCh) from baseline (time 0) and from the preceding time points were calculated for all time points (time i) by the formula presented below:

$$PCH_i = \left( \frac{X_i - X_0}{X_0} \right)$$

The distribution of the percent change was normal based on normal probability plot. Means and 95% confidence intervals of percent changes were estimated. Prediction linear regression models were built based on log-transformed enzyme levels for each time point. RMSE (root mean squared error) and R-squared were calculated using 10-fold cross-validation to explore the performance of the models. The underlying assumption for linear regression was tested and validated. The correction factors were calculated for each time point in order to correct the decline in level of enzymes from the baseline using the following formula:

$$\text{Ln}(X_i) = \alpha + \beta \text{Ln}(X_0)$$

Statistical analyses were performed using Stata software (version 12.0; StatCorp).

Results
Principal Findings
Table 1 presents the results of the median (25th (Q1) and 75th (Q3) percentiles) of liver serum enzymes at different time points collected from 400 recruited participants (mean age: 47.4 ± 15.9 years), including 253 (63.0%) females. The results show a significant declining trend in ALT, AST, and ALP at all points compared to baseline. The exception was GGT that did not show a significant decline three months after the baseline. The p for trend in GGT was not significant, either.

Table 2 and Figure 1 demonstrate the percent changes from baseline, as well as trimonthly percent change in each time point compared to the preceding point of time. ALT shows the most pronounced decline within three months after the baseline (-38.46%; 95% CI -40.3 to -36.6). The decline becomes less steep after three months (-5.68; 95%
However, for AST, the trimonthly percent changes were significantly negative at all time points compared with their preceding points of time.

AST and ALP showed minimal yet significant decline at three months compared to baseline (-2.44% and -0.80% respectively) and almost non-significant declines after six months compared to the preceding three months (-1.42% and -6.51%). However, there was a significant decline in AST and ALP between six months and two years after the baseline (-10.13% and -7.12% respectively), which levels off after two years (-2.58% and 0.52% respectively).

As for GGT, the percent declines were quite minimal at all points of time and the final percent change from baseline up to five years afterwards was only -3.30% (-6.0 to -0.5). Of note, ALP and GGT concentrations were not altered statistically until three months, but decreased significantly afterwards.

Modeling for the Time-Dependent of Analytes Concentrations
Considering the systematic decline in concentration of liver enzymes (ALT, AST, and ALP) from baseline, we tried to develop a prediction model in order to correct for this decline. We used a linear regression analysis on
Stability of Liver Enzymes in Blood Serum

Many studies have focused on the effect of short-time storage on some of liver function enzymes\textsuperscript{16-23} or the effect of temperature, but few addressed long-term changes.\textsuperscript{6,24,25} The only study that has the greatest similarity in temperature and time of storage with ours was conducted by Yu et al.\textsuperscript{25} In this study, investigators examined the stability of chronic hepatitis-related parameters including GGT and ALT in serum samples up to 10 years of storage. They demonstrated that temperatures of \(-70^\circ\text{C}-80^\circ\text{C}\) were necessary to store serum samples for a long time. On the other hand, storing samples at \(-20^\circ\text{C}\) is only suitable for short-term analysis. They found that GGT concentrations do not change after storage for less than three years; however, they change significantly even at \(-70^\circ\text{C}\) after 7 and 10 years of storage. In the study, ALT concentrations showed a significant decrease after one year of storage at either \(-20^\circ\text{C}\) or \(-70^\circ\text{C}\) from the baseline. Our study also shows no significant difference for GGT after five years and a decrease in ALT concentrations until one year of storage at \(-70^\circ\text{C}\) which is in agreement with that study.\textsuperscript{25} On the contrary, Kachhawa et al. reported that the activity of liver enzymes (AST, ALT, and ALP) significantly decreased after 30 days of storage at \(-20^\circ\text{C}\) and it is not recommended to store samples at this temperature for more than one month.\textsuperscript{26}

A few studies showed an increase in serum samples after long time storage. For example, in a study conducted on the serum samples of ten adult males in India,\textsuperscript{6} they found an increase in AST and ALT concentrations after a short time of 72 hours and a long time of three months. These

log-transformed levels of enzymes. The results of the model are presented in Table 3. As can be observed in this table, the performance of prediction models decreases (in terms of RMSE and R-squared) as the time period since baseline increases. The performance is acceptable until one year after the baseline collection of the samples, as R-squared remains above 0.7. Yet, the performance of the prediction model declines substantially two and five years after the baseline. The exception was GGT. As we did not observe any significant systematic change in its levels from baseline, we did not develop any prediction model for this liver enzyme.

Figures of trend demonstrate the gradual decrease in ALT, AST and ALP values over the time period, but GGT indicated a small increase three years after initial sample collection.

**Discussion**

In this study, we explored the stability of liver enzymes after long-term storage. Our main findings reveal a significant declining trend for ALT, AST, and ALP at all points after the initial storage compared to baseline. GGT was an exception. ALT shows the steepest decline within three months after storage but the decline in AST was more gradual. The decline in AST and ALP was significant between six months and two years after the baseline. The concentration of ALP and GGT did not change statistically until three months, but decreased significantly afterwards. We developed a prediction model in order to correct for this decline, which had an acceptable performance.

![Image](image-url)
Table 3. Regression Coefficients and Models Performance for Prediction Linear Models

<table>
<thead>
<tr>
<th></th>
<th>ALT</th>
<th>AST</th>
<th>ALP</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 months</td>
<td>α</td>
<td>β</td>
<td>R-squared</td>
<td>RMSE</td>
</tr>
<tr>
<td>a</td>
<td>0.72 (0.68 to 0.76)</td>
<td>0.87 (0.83 to 0.9)</td>
<td>0.91 (0.88 to 0.94)</td>
<td>0</td>
</tr>
<tr>
<td>β</td>
<td>1.22 (1.12 to 1.32)</td>
<td>0.42 (0.31 to 0.53)</td>
<td>0.49 (0.34 to 0.63)</td>
<td>0</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.29</td>
<td>0.14</td>
<td>0.09</td>
<td>0</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.76</td>
<td>0.82</td>
<td>0.92</td>
<td>0</td>
</tr>
<tr>
<td>6 months</td>
<td>α</td>
<td>β</td>
<td>R-squared</td>
<td>RMSE</td>
</tr>
<tr>
<td>a</td>
<td>0.75 (0.71 to 0.79)</td>
<td>0.88 (0.84 to 0.92)</td>
<td>0.86 (0.83 to 0.89)</td>
<td>0</td>
</tr>
<tr>
<td>β</td>
<td>1.17 (1.08 to 1.27)</td>
<td>0.39 (0.28 to 0.5)</td>
<td>0.82 (0.66 to 0.97)</td>
<td>0</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.28</td>
<td>0.14</td>
<td>0.10</td>
<td>0</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.78</td>
<td>0.83</td>
<td>0.89</td>
<td>0</td>
</tr>
<tr>
<td>1 year</td>
<td>α</td>
<td>β</td>
<td>R-squared</td>
<td>RMSE</td>
</tr>
<tr>
<td>a</td>
<td>0.76 (0.71 to 0.81)</td>
<td>0.79 (0.74 to 0.83)</td>
<td>0.89 (0.86 to 0.92)</td>
<td>0</td>
</tr>
<tr>
<td>β</td>
<td>1.47 (1.37 to 1.57)</td>
<td>0.76 (0.63 to 0.89)</td>
<td>0.69 (0.53 to 0.84)</td>
<td>0</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.32</td>
<td>0.18</td>
<td>0.10</td>
<td>0</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.73</td>
<td>0.72</td>
<td>0.89</td>
<td>0</td>
</tr>
<tr>
<td>2 years</td>
<td>α</td>
<td>β</td>
<td>R-squared</td>
<td>RMSE</td>
</tr>
<tr>
<td>a</td>
<td>0.48 (0.43 to 0.53)</td>
<td>0.56 (0.51 to 0.61)</td>
<td>0.62 (0.58 to 0.66)</td>
<td>0</td>
</tr>
<tr>
<td>β</td>
<td>2.4 (2.33 to 2.48)</td>
<td>1.68 (1.57 to 1.79)</td>
<td>2.40 (2.20 to 2.61)</td>
<td>0</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.42</td>
<td>0.23</td>
<td>0.18</td>
<td>0</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.51</td>
<td>0.54</td>
<td>0.68</td>
<td>0</td>
</tr>
<tr>
<td>5 years</td>
<td>α</td>
<td>β</td>
<td>R-squared</td>
<td>RMSE</td>
</tr>
<tr>
<td>a</td>
<td>0.33 (0.27 to 0.39)</td>
<td>0.35 (0.3 to 0.4)</td>
<td>0.2 (0.88 to 0.95)</td>
<td>0</td>
</tr>
<tr>
<td>β</td>
<td>2.77 (2.71 to 2.84)</td>
<td>2.3 (2.2 to 2.39)</td>
<td>0.82 (0.67 to 0.97)</td>
<td>0</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.49</td>
<td>0.27</td>
<td>0.17</td>
<td>0</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.32</td>
<td>0.40</td>
<td>0.69</td>
<td>0</td>
</tr>
</tbody>
</table>

ALT, alanine aminotransferase; AST, aspartate aminotransferase; ALP, alkaline phosphatase; RMSE, root mean squared error.

results could be due to hemolysis or increase in the activity of enzymes after storage. In our study, all enzymes showed a decrease over time but GGT demonstrated a small increase three years after collection.

In one study,27 conducted on long term stability of cancer biomarkers, serum samples of 34 volunteers were prepared within two hours and stored at -20°C temperature. After two days, samples were divided and stored at -20°C and -80°C until the days of analysis. After one year, the amount of ALT decreased by 40% at -20°C but showed a good concentration at -80°C, which is in contrast with our study. We showed a deep decline within three months of storage. This difference could be explained by the different temperature of the studies. GGT concentration showed no statically significant difference between these two temperatures after 12 months and also no significant difference from baseline.

The results of our study showed that it is preferable to measure liver enzymes, especially ALT, within three months after baseline collection, as a large proportion of the decline occurs during this period. Yet serum concentrations of ALT continue to decline consistently until two years and level off afterwards.

Our study also showed that due to the high proportion of samples in which liver enzymes cannot be detected after two years of storage, it is not advised to store samples for more than two years. Although this result is not the same for all the enzymes. As for GGT, we did not see any significant changes even after five years; so, we could not discuss the model for this enzyme.

The prediction models had satisfactory performance until one year after the baseline collection of samples but the models were inaccurate after one year.

To the best of our knowledge, this study is the first in which all four liver enzymes were investigated for a long period of five years. Additionally, we used a large sample of 400 randomly selected participants of both sexes in comparison with small sample sizes in many other studies.6,27 Many studies kept their samples at -20°C but we stored our samples at -70°C which is the standard temperature for long term storage.24,25,28

The main limitation of our study is the lack of temperature monitoring of the freezers. In our study, we tried to limit the effect of hemolysis by visual examination, but as shown in the study by Koseoglu et al.,29 even in macroscopically invisible hemolysis, AST showed an increase in level.

It is recommended to study the effect of long-time storage at different temperatures and other types of analytes. The effect of age and sex on the concentration of analytes over time can be explored.

The key finding of this article is that, even at -70°C, long-term storage of serum samples significantly decreases the concentration of the liver function enzymes from baseline, except for GGT. It is not recommended to store samples for more than two years as liver enzymes are not detectable afterwards. We also suggest using the prediction models that have satisfactory performance until one year.
after sample collection. These findings can be considered in studies that are based on long-term storage of biosamples.

Authors’ Contribution
SK conceived, designed and supervised the study and edited the manuscript. SGS designed, conducted initial analysis and edited the manuscript. ES drafted the manuscript. MS contributed in drafting tables and initial data analysis. SK, SGS and MA provided experts’ opinion and commented on intellectual development of manuscript. All authors read and approved the final draft. RM, SM and HP are guarantors.

Conflict of Interest Disclosures
None declared.

Ethical Statement
The research protocol was approved and the use of serum samples was authorized by the Institutional Review Board and Ethics Committee of the DDRI of TUMS.

References