Determining Stability of Stored Samples Retrospectively: The Validation of Glycated Albumin

David M. Nathan,1* Michael W. Steffes,2 Wanjie Sun,3 Gregory P. Rynders,2 and John M. Lachin3

BACKGROUND: Determining the stability of stored samples for assays that were not available at the time of original collection is problematic. To assess sample stability for a relatively new assay of glycated albumin (GA), we first measured GA in fresh samples and in samples stored for 19–23 years. We then compared the regression of the contemporaneous glycohemoglobin (Hb A1c) values against the GA results from fresh vs stored samples, reasoning that similar slopes and intercepts would provide strong, albeit indirect, support for the stability of the stored samples for GA measurements.

METHODS: We assayed 90 samples frozen for 19–23 years and 90 fresh samples from participants in the Diabetes Control and Complications trial cohort for GA. Hb A1c was measured contemporaneously in fresh samples at each time period. A single normal-errors linear model regressed the Hb A1c values on the GA, with an additional effect for collection period (fresh vs stored for GA) and the interaction of period and GA.

RESULTS: Analysis of the regressions lines between GA and Hb A1c revealed intercepts (3.69 and 2.97 for the fresh and stored samples, respectively) and slopes (0.198 vs 0.187) that were not significantly different (P = 0.182 and P = 0.639, respectively).

CONCLUSIONS: This simple approach can be used to assess the stability of stored samples in new assays. Samples stored for as long as 23 years are suitable for the GA assay.

© 2010 American Association for Clinical Chemistry

The traditional approach to determine the stability of an assay over time in stored specimens is to conduct the assay on a split aliquot from a collection of fresh specimens and then again on a split aliquot after an extended period of freezing. If the mean and variance of the fresh and stored specimens are the same, and there is a strong linear correlation, then the stability of samples used in the assay is established.

Long-term, longitudinal clinical studies often store biological samples over time with the aim of performing assays that become available during the course of the study. Most relatively new assays have limited data with regard to sample stability, and almost no assays have stability data that extend over the time period of very-long-term studies. The Diabetes Control and Complications Trial (DCCT)4 (1) and its long-term follow-up study, the Epidemiology of Diabetes Interventions and Complications (EDIC) (2), have been following a cohort of 1441 patients enrolled in 1983–1989 with periodic collections of biological specimens for assay of biochemical and metabolic measurements, and with long-term storage of split aliquots. Levels of glycemia have been assessed continuously with a glycohemoglobin (Hb A1c) assay that has been stable over the 25-year period of the study (3).

Recently, a DCCT/EDIC ancillary study was proposed that would use a relatively new glycated albumin assay (4) on samples, some stored for as long as 25 years. Hb A1c, expressed as the percentage of total hemoglobin, reflects the degree of glycation of hemoglobin in the intracellular space over the preceding 90–120 days. Thus, it is a poor marker of day-to-day variability in levels of glycemia. Conversely, glycated albumin, expressed as the percentage of total albumin, reflects the degree of glycation over the preceding 2–4 weeks, and thus reflects shorter-term changes in glycemia and, potentially, glycemic variability better than the Hb A1c assay. Previous analyses have shown a strong association of Hb A1c values over the 25 years of DCCT and EDIC with the risks of micro- and macrovascular complications in type 1 diabetes (5, 6). The ancillary study proposed to examine the association of glycated albumin with these outcomes and determine whether it was an independent risk factor in addition to Hb A1c.

1 The Diabetes Center, Massachusetts General Hospital, and Harvard Medical School, Boston, MA; 2 University of Minnesota Medical School, Minneapolis, MN; 3 Biostatistics Center, George Washington University, Rockville, MD.
* Address correspondence to this author at: MGH Diabetes Center, 50 Staniford St., Suite 340, Boston, MA 02114. Fax 617-726-6781; e-mail dnathan@partners.org.

Trial registration: clinicaltrials.gov, identifiers NCT00360815 and NCT00360893. Received May 4, 2010; accepted August 9, 2010. Previously published online at DOI: 10.1373/clinchem.2010.150250

4 Nonstandard abbreviations: DCCT, Diabetes Control and Complications Trial; EDIC, Epidemiology of Diabetes Interventions and Complications; Hb A1c, glycohemoglobin.
Sample Stability for Glycated Albumin Assay

To do so, it was necessary to determine the stability of this new assay when performed on samples in storage for an extended period. Because there was no direct way to compare the results of the assay performed on freshly collected samples and again later on a split aliquot after extended storage, it was necessary that we develop an indirect, retrospective method of assessing the stability of the glycated albumin assay in long-term stored samples.

Herein, we describe a method to perform retrospective validation of the stability of stored samples for new assays and its application to the glycated albumin assay.

Methods

STUDY PATIENTS

The DCCT was a controlled clinical trial conducted between 1983 and 1993 to compare the effects of intensive therapy, aimed at achieving near-normal glycemia, with conventional (at the time) diabetes therapy on the long-term complications of type 1 diabetes mellitus (1). The study enrolled 1441 volunteers with type 1 diabetes and, after 6.5 years of mean follow-up, demonstrated conclusively that intensive therapy had a major and consistent salutary effect on the prevention and delay of diabetic complications (1). The major assessment of glycemia during DCCT was the Hb A1c assay, measured every 3 months. The HPLC method used has been described in detail, as has the long-term stability of the assay (3). After the end of the DCCT in 1993, the DCCT volunteers were invited to join an observational, long-term follow-up, the EDIC study. Ninety-six percent of the surviving members of the DCCT cohort joined EDIC, and they have been followed to the present day with methods that were similar, if not identical, to those used during the DCCT (7). In particular, Hb A1c assays are performed with the same methods as in the DCCT, but now annually (3, 8).

ASSAYS

The glycated albumin assay that we evaluated was the Lucica GA-L™ kit (Asahi Kasei Pharma Corporation), an enzymatic method coupled to a colorimetric output. We performed the assay according to the manufacturer’s instructions on a Modular P analyzer™ (Roche) and as published in the literature (9). We measured concentrations of albumin and glycated albumin and calculated percent glycated albumin. We established imprecision of the assays by performing repeated assays on fresh plasma samples: the intra- and interassay CVs were <2% for low and high standard samples with a coefficient of reliability >0.99 from split duplicate samples; i.e., <1% of total variation is attributable to the sources of random variation (for example, within or between batches).

SAMPLES

Whole blood samples in EDTA-containing collection tubes from research participants at the 27 DCCT/EDIC clinical centers during their regularly scheduled visits were shipped on ice overnight to the DCCT/EDIC Central Biochemistry Laboratory at the University of Minnesota. Samples were stored at −70°C until being thawed for assay. After processing, plasma samples were similarly stored.

STUDY DESIGN

Although Hb A1c and glycated albumin are both the products of glycation, owing to the differences described above, a perfect linear relationship between the 2, or an $R^2$ of 1.0, was not expected. Rather, a substantial correlation was expected, with some scatter about the line of regression. If measured glycated albumin values were stable in stored specimens, then the nature of the regression between the glycated albumin from a stored specimen and the Hb A1c measured from the original specimen on receipt should be similar to the regression relationship between the glycated albumin and the Hb A1c when both are measured from recent collections. This hypothesis was testable because the Hb A1c assay over the 25 years of the DCCT/EDIC has been rigorously controlled internally, its precision determined by assays on external masked split-duplicate collections. For the statistical model on which this approach is based, see the Appendix in the Data Supplement that accompanies the online version of this article at http://www.clinchem.org/cgi/content/vol57/issue2.

To cover a wide range of Hb A1c values, we randomly selected 18 frozen specimens from each quintile of the Hb A1c concentrations determined at the time of the original Hb A1c assay (total of 90 samples). All the frozen specimens were collected during 1986–1990 from a visit 1 year after randomization into the DCCT. They had been stored frozen at −70°C for 19–23 years. The samples were thawed and glycated albumin assayed from the stored specimen. We also selected 90 specimens collected during 2009, shipped frozen to the central laboratory, then thawed and assayed for glycated albumin and compared to Hb A1c values from whole blood collected at the same time.

STATISTICAL METHODS

A single normal-errors linear model regressed the Hb A1c values on the glycated albumin with an additional
effect for collection period (fresh vs stored), and the interaction of period and glycated albumin. In effect, the model fitted separate regressions of Hb A1c on glycated albumin within a period and allowed a test of the difference between the 2 regression equations. This also provided the SE of the intercept and slope and the corresponding 95% CI for each period. We verified the assumptions of normally distributed errors and homoscedastic variances by statistical testing.

For a regression of $Y$ on $X$, with slope $\beta$ and means $\mu_Y$ and $\mu_X$, the intercept is $\alpha = \mu_Y - \beta \mu_X$, which in turn depends on the means of the 2 variables. Thus the intercept in each period is a function of the slope and the mean levels of Hb A1c and glycated albumin in samples from each period. If the ratio of the variance of Hb A1c to that of glycated albumin is the same for the stored and fresh samples, then the difference in slopes is proportional to the difference in correlations. It was thought that the $R^2$ of the regression could be about 0.6, or a correlation of 0.77, in the fresh EDIC samples. A sample size of 90 per group then provided 80% power to detect a 17% reduction in this correlation in the stored DCCT samples; equivalently, a 17% reduction in the slope.

Results

The biochemical characteristics of the 90 fresh and 90 stored samples are shown in Table 1. The distributions of glycated albumin and Hb A1c during each period were close to normal, with a slight right skewness.

The relationships between Hb A1c, measured in the original specimens and the glycated albumin levels from the fresh and stored samples are shown in Fig. 1. The relationships were well described as linear regressions, with an $R^2$ value of 0.592 for the fresh samples and 0.696 for the stored samples (both $P < 0.0001$). The intercepts (3.69 and 2.97 for the fresh and stored samples, respectively) were not significantly different ($P = 0.58$), which was also the case for the slopes (0.198 vs 0.187, $P = 0.639$). The residuals from the above regressions are approximately normally distributed, apart from 1 outlier (noted below), and the variances appear to be homoscedastic, i.e., constant across the range of the assay.

The regression lines in Fig. 1 for the recent and stored specimens are almost perfectly parallel, with nearly identical slopes. However, the 2 lines are slightly offset due to the difference in the intercepts. As described in Methods, the small difference in intercepts for recent vs stored specimens results from significantly higher levels of glycated albumin in stored vs recent specimens. Given the near equality of slopes, there would also be a near equality of intercepts if the mean level of glycated albumin and Hb A1c were the same in the 2 periods.

There is 1 outlier among the recent specimens, with a glycated albumin of about 15% and an Hb A1c of about 10%, that is possibly a result of errors in handling or labeling. Removing this observation has minimal effect on the results: the intercept decreases slightly to 3.3, still greater than in the stored specimens; the slope increases to 0.199, nearly identical to the stored specimens; and the $R^2$ increases from 0.592 to 0.673. However, the mean glycated albumin value remains lower than that in the stored specimens.

Discussion

Establishing the stability of samples in storage is conventionally accomplished by assaying samples at the time they are collected and repeating the assay on split duplicates over time. For example, Selvin et al. recently used this approach to establish that the Hb A1c assay was stable in samples frozen for 11–14 years (10). For newly developed assays, however, establishing stability

<table>
<thead>
<tr>
<th>Table 1. Biochemical characteristics of stored samples from 1986–1990 compared with recent samples from 2009.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1986–90</td>
</tr>
<tr>
<td>n</td>
</tr>
<tr>
<td>Hb A1c, %</td>
</tr>
<tr>
<td>Glycated albumin, %</td>
</tr>
<tr>
<td>Regression relationship</td>
</tr>
<tr>
<td>Intercept</td>
</tr>
<tr>
<td>Slope</td>
</tr>
<tr>
<td>$R^2$</td>
</tr>
</tbody>
</table>

* Data are mean (SD) unless noted otherwise.
for stored samples remains a challenge. Herein we describe a simple approach that can be used to do so retrospectively. All that is needed is to identify a measurement (such as the Hb A1c) that has been assayed longitudinally over the length of the study from fresh collections and that is expected to have a moderate correlation with the new assay. Then, if the slope of the regression of the longitudinal measurement on the new assay obtained from fresh specimens is the same as that obtained from long-term stored specimens, one can conclude that the samples frozen and measured in the new assay are stable over time.

The glycated albumin values from recent collections are significantly lower than those from stored specimens, the difference being significant by a Wilcoxon rank sum test. The difference between the fresh and frozen stored samples could represent an effect of prolonged freezing. However, it would have been expected that the stored values might be lower as a result of freezing, not higher. Rather than related to freezing, the difference in values over time could simply reflect the difference between the level of glucose control during the 2 periods. The Hb A1c is greater by 0.24% in the 1986–1990 samples than in the recent samples, reflecting the higher range of values among subjects treated conventionally during the DCCT. After the DCCT demonstrated the benefits of intensive therapy in 1993, all subjects were trained in the implementation of intensive therapy, and the Hb A1c values in the former conventional group dropped by approximately 1%. In the intensive treatment group, they rose by approximately 0.8%. Further, long-term storage would have been expected to dilute the estimate of the slope of the regression of Hb A1c on glycated albumin (the Hb A1c values obtained from the original collections). However, both the slope estimate and the associated $R^2$ were somewhat greater from the stored collections than the fresh collections.

In practice, if the slope of the relationship between the 2 variables, 1 measured longitudinally in real time (e.g., Hb A1c), the other from fresh vs frozen specimens (glycated albumin), is the same for the 2 periods, then it matters little which of the 2 is the independent variable. In our case, we chose to regress the Hb A1c on the glycated albumin because in future analyses we wished to assess the fraction of the Hb A1c effect on outcomes explained by the glycated albumin. Further, as shown in the Appendix in the online Data Supplement, if there is instability in the glycated albumin assay, then the effects on the slope will be more apparent when Hb A1c is regressed on the glycated albumin than vice versa.

We are not aware that the indirect approach employed herein, with assay of 1 marker (Hb A1c) at the time of collection and assay of another from stored frozen specimens and fresh specimens, has been applied to the evaluation of the long-term stability of

---

**Fig. 1.** Linear regression of glycated hemoglobin on glycated albumin values for recent (2009, 90 samples, red symbols and line) and stored (1986–1990, 90 samples, black symbols and line) samples.

The linear regressions had an $R^2$ value of 0.592 for the recent samples and 0.696 for the stored samples (both $P < 0.0001$). The intercepts (3.69 and 2.97 for the fresh and stored samples, respectively) were not significantly different ($P = 0.182$), which was also the case for the slopes (0.198 vs 0.187, $P = 0.639$).
other biomarkers. However, a similar approach was recently used to assess the stability of 25-OH vitamin D levels in stored specimens (11). Bodnar et al. were able to detect seasonal and racial differences in 25-OH vitamin D concentrations in specimens frozen for 40 years that paralleled the results of recently collected specimens. Although the concentrations in the frozen specimens were lower than those in the recent specimens, the coefficients for the racial and seasonal effects were not significantly different between the stored and recent collections.

Although homogeneous regression slopes for fresh and frozen specimens may be a strong indication of stability, the converse is not necessarily the case. Results could be affected if the regression relationship depended on other factors that might also differ among the fresh vs stored specimens, such as treatment, age, or sex. It would be preferable that the comparison marker (like Hb A1c herein) be known not to be affected by such other factors and that the assumption of a constant correlation and slope over time is plausible for the measures involved.

In conclusion, we describe a simple approach that can be used to assess retrospectively the stability of stored samples in a new assay. Essentially, the new assay is benchmarked against another assay that has been measured longitudinally from fresh collections over the span of the study. When the longitudinally measured Hb A1c was regressed on glycated albumin assayed on stored vs fresh specimens, near-identical slopes were observed. This strongly demonstrates the stability of the glycated albumin assay in long-term stored specimens. This approach could also be applied to establish the stability of samples in long-term storage for other assays.

References