BACKGROUND: Because cerebrospinal fluid (CSF) is in close contact with diseased areas in neurological disorders, it is an important source of material in the search for molecular biomarkers. However, sample handling for CSF collected from patients in a clinical setting might not always be adequate for use in proteomics and metabolomics studies.

METHODS: We left CSF for 0, 30, and 120 min at room temperature immediately after sample collection and centrifugation/removal of cells. After trypic digestion at 2 laboratories by nanoLC Orbitrap-MS and chipLC QTOF-MS, CSF proteomes were analyzed. Metabolome analysis was performed at 3 laboratories by NMR, GC-MS, and LC-MS. Targeted analyses of cystatin C and albumin were performed by LC-MS/MS in the selected reaction monitoring mode.

RESULTS: We did not find significant changes in the measured proteome and metabolome of CSF stored at room temperature after centrifugation, except for 2 peptides and 1 metabolite, 2,3,4-trihydrobutanoic acid, of 5780 identified peptides and 93 identified metabolites. A sensitive protein stability marker, cystatin C, was not affected.

CONCLUSIONS: The measured proteome and metabolome of centrifuged, human CSF is stable at room temperature for up to 2 hours. We cannot exclude, however, that changes undetectable with our current methodology, such as denaturation or proteolysis, might occur due to sample handling conditions. The stability we observed gives laboratory personnel at the collection site sufficient time to aliquot samples before freezing and storage at $-80^\circ$C.

Conditions between the time of collection of a biological sample and its arrival at the analytical research laboratory might not always be adequate for subsequent proteomics and metabolomics analyses, especially in cases where the sample collection was not originally performed with these large-scale analyses in mind. To detect reliable molecular biomarkers, it is imperative to handle biological fluids according to standardized procedures and to evaluate the effects of preanalytical parameters on the final result to avoid artifacts (1, 2). Earlier studies on urine, plasma, and cerebrospinal fluid (CSF)9 have shown that sample handling can affect the stability of proteins as well as metabolites (3–11). Standardized sample handling is also important when trying to compare results between different laboratories (12–14). In the search for molecular biomarkers related to disorders of the central nervous system, CSF is the most promising biofluid because of its close contact to the affected tissue (13, 15–20).

In this study, we analyzed a set of human CSF samples to assess protein and metabolite stability at room temperature after a low-speed centrifugation step to remove cells. To cover a wide range of proteins and metabolites, we combined the results from several analytical platforms, including LC-MS, GC-MS, and nuclear magnetic resonance (NMR).
Materials and Methods

SAMPLE SET
We obtained 6 human CSF samples from the Department of Neurology at the Erasmus University Medical Center (Rotterdam, The Netherlands). The CSF samples were collected as part of routine clinical examination of patients with various symptoms (Table 1). All samples were collected via lumbar puncture between the third and fourth lumbar vertebrae with a SpinoCan needle (0.90 × 88 mm). The Medical Ethical Committee of the Erasmus University Medical Center approved the study protocol, and all patients gave their informed consent. Samples were centrifuged (10 min at 956 × g) within 5 min after collection to remove cells. Aliquots were directly snap-frozen in liquid nitrogen or left at room temperature for 30 and 120 min before snap-freezing and storage at −80 °C. Routine CSF diagnostics, including total protein and albumin concentration measurements as well as intrathecal cell count, were performed. To eliminate the possibility that samples were contaminated with blood, hemoglobin and apolipoprotein B100 were measured and confirmed to be absent in all specimens. Sample H1 was analyzed by use of the chipLC quadrupole time-of-flight (QTOF) MS and the nanoLC Orbitrap tandem mass spectrometry (MS/MS) platforms only. Samples H2–H6 were analyzed by all platforms. We performed protein digestion for proteomics analysis as described by all platforms. We performed protein digestion for pro-

<table>
<thead>
<tr>
<th>Sample</th>
<th>Age, years</th>
<th>Sex</th>
<th>Diagnosis</th>
<th>Protein concentration, mg/L</th>
<th>Clinic albumin concentration, mg/L</th>
<th>SRM albumin concentration, mg/L*</th>
<th>Cells/μL after centrifugation, n</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1b</td>
<td>49</td>
<td>M</td>
<td>Migraine</td>
<td>415</td>
<td>193</td>
<td>192.3</td>
<td>0</td>
</tr>
<tr>
<td>H2</td>
<td>56</td>
<td>M</td>
<td>Idiopathic intracranial hypertension</td>
<td>472</td>
<td>247</td>
<td>237.5</td>
<td>0</td>
</tr>
<tr>
<td>H3</td>
<td>69</td>
<td>F</td>
<td>Headache</td>
<td>395</td>
<td>236</td>
<td>221.0</td>
<td>0</td>
</tr>
<tr>
<td>H4</td>
<td>48</td>
<td>M</td>
<td>Idiopathic intracranial hypertension</td>
<td>436</td>
<td>241</td>
<td>225.6</td>
<td>0</td>
</tr>
<tr>
<td>H5</td>
<td>29</td>
<td>F</td>
<td>Clinical isolated syndrome (neuromyelitis optica)</td>
<td>387</td>
<td>226</td>
<td>213.9</td>
<td>0</td>
</tr>
<tr>
<td>H6</td>
<td>38</td>
<td>F</td>
<td>Epilepsy</td>
<td>381</td>
<td>184</td>
<td>194.2</td>
<td>0</td>
</tr>
</tbody>
</table>

* Average over 3 time points (online Supplementary Table S4).

Sample H1 was analyzed only with respect to proteomics.

ChipLC QTOF-MS PROTEOMICS ANALYSIS
In quintuplicate, we randomly injected 0.5 μL trypsin-digested CSF (H1–H6: T0, T30, T120), and we injected 0.5 μL digested QC samples (pooled CSF spiked with cytochrome C; Fluka, part #30396; final concentration, 375 fmol/μL) and blanks between every tenth sample for LC-MS analysis on an Agilent chipLC QTOF-MS system as reported (21). Enrichment and separation was done using an LC chip (G4240-63001 SPQ110, Agilent Technologies) (separation column, 150 mm × 75 μm Zorbax 300SB-C18, 5 μm; trap column, 160-nL Zorbax 300SB-C18, 5 μm). The LC separations were carried out as described using the following gradient: 80-min linear gradient from 3% to 40% B (0.1% formic acid in acetonitrile) in A (0.1% formic acid in ultrapure water), followed by 10-min linear gradient from 40% to 50% B and 10-min linear gradient from 50% to 3% B (21). We performed MS analysis under the following conditions: mass range, 200–2000 m/z in profile mode; acquisition rate, 1 spectrum/s; fragmenter voltage, 175 V; skimmer voltage, 65 V; OCT 1 RF Vpp, 750 V. The spray voltage was approximately 1800 V, and the drying gas (N2) was 6 L/min at a temperature of 325 °C. Each spectrum was mass corrected by use of internal standards (methyl stearate, m/z 299.294457, and HP-1221, m/z 1221.990637) evaporating from a wetted wick inside the spray chamber. Reproducibility was monitored on selected cytochrome C peaks in the QC samples; mass difference between theoretical and the measured values was within 4 ppm. The selected peaks showed a peak area CV of <20% and a retention time CV of <2%.

We processed the data using a pipeline developed in C++ as described (21, 22). MzData.XML data were converted to ASCII format over a mass range of 200–1600 m/z (outside this range no multiply charged ions were detected), a retention time range of 3–80 min (peptide elution range), and an intensity threshold of 300 counts. We applied a double cross-validated nearest-shrunken centroid (NSC) algorithm to the complete peak matrix; for a certain shrinkage value, the NSC comparison gives a cross-validation error between 0 and 1, where 1 implies that both classes are assigned to the wrong class, 0.5 is a random class assignation, and 0 means that both classes are correctly assigned (23). NSC-selected features were compared by univariate statistical analysis (Student t-test with Bon
ferroni correction for multiple comparisons) and ANOVA (Microsoft Excel 2007 and SPSS 16.0). The features were considered significantly different with a P value <0.05 (T0 vs T120 and T0 vs T30) in at least 5 of 6 samples (T0 vs T120 and T0 vs T30). Each discriminatory feature was analyzed by targeted tandem MS for identification. We applied principal component analysis (PCA) (24) to the complete peak matrix (10 000 peaks) as well as to the NSC-selected features (MatLab, R2009a).  

For visualization, we created box-and-whisker plots in Origin 7.0.

**NANO LC ORBITRAP-MSMS SHOTGUN PROTEOMICS ANALYSIS**  
Trypsin-digested CSF samples (H1–H6: T0, T30, T120) were injected in random order and analyzed by MS/MS (shotgun approach) on an Ultimate 3000 nano LC system (Dionex) online coupled to a hybrid linear ion trap/Orbitrap mass spectrometer (LTQ Orbitrap XL; Thermo Fisher Scientific) as described (21).

Data files were analyzed and preprocessed using the Progenesis LC-MS software package (Nonlinear Dynamics). Retention times were aligned, and the intensities of the ions were normalized based on the total ion current. To assess interpatient variability, all identified peaks were analyzed by PCA. All identified peaks were also analyzed by the NSC algorithm for classification (23). Peptides were analyzed for differential abundance between the groups by ANOVA. P values <0.01 were considered significant.

We searched all MS/MS spectra against the Uniprot/SwissProt database (version 57.6, taxonomy Homo sapiens, 20 070 sequences) using Mascot (version 2.2.06). Search parameters were as follows: parent ion tolerance, 2 ppm; amino acid modifications, carbamidomethylation of cysteine (fixed) and oxidation of methionine (variable).

**NANO LC-MS/MS ANALYSES IN THE SELECTED REACTION MONITORING MODE**  
Trypsin-digested CSF samples (H1–H6: T0, T30, T120) were spiked with known concentrations of stable isotope-labeled peptide standards corresponding to 1 peptide (FQNALLVR) of human serum albumin and 2 (ALDFAVGEYKN) of human cystatin C for quantification by selected reaction monitoring (SRM) (Supplementary Tables S1 and S2, which accompany the online version of this article at http://www.clinchem.org/content/vol57/issue12). We performed chromatographic separation of spiked CSF digests on an Ultimate 3000 nano LC system (Dionex). Spiked CSF digest (1 μL) was loaded onto a C18 trap column (PepMap C18, 300 μm ID by 5 mm, 5 μm particle size and 100 Å pore size; Dionex) and washed for 5 min at a flow rate of 20 μL/min 0.1% TFA in H2O. Next, the trap column was switched in line with the analytic column (PepMap C18, 75 μm ID by 150 mm, 3 μm particle size and 100 Å pore size; Dionex). Peptides were eluted at a flow rate of 300 nL/min with the following gradient: 0%–45% solvent B in 30 min, where solvent A is H2O/acetonitrile (ACN) 98%/2% (vol/vol), 0.1% formic acid (FA) and solvent B is H2O/ACN 20%/80% (vol/vol), 0.1% FA. The separation of the peptides was monitored with a UV detector (absorption at 214 nm).

MRM detection was performed by means of a triple quadrupole tandem mass spectrometer (4000 QTRAP; AB Sciex) in the positive ion mode. As technical control for the measurements, a single spiked CSF digest was measured after every sixth run. A technical control for the enzymatic digestion (1 sample digested at 3 separate times) was also included in this quantitative analysis. We performed data analysis using the MRM data analysis program Skyline (version 0.7) (25), with the ratio of the analyte peptide to the known concentration of the spiked isotope-labeled internal peptide standard to calculate the concentrations of the original peptides. For the cystatin C peptide, we used the average of both fragment ion ratios for the determination of the protein concentration, and we used a paired, 2-sided t-test to test for differences in peptide concentrations between the time points.

**GC-MS METABOLOMICS ANALYSIS**  
CSF samples (H2–H6: T0, T30, T120) were treated with an oximation reagent followed by silylation before GC-MS analysis (21, 26). Each sample was injected twice in random order and analyzed on an Agilent 6890 gas chromatograph coupled to an Agilent 5973 quadrupole mass spectrometer as described (21).

Peaks were characterized by retention time and m/z ratio and identified by comparison with a spectral database (TNO) (21). We analyzed all detected metabolites by PCA. A 2-tailed Student t-test was applied to all known metabolites (T0 vs T30 and T0 vs T120). Metabolites with a P value <0.05 were considered to be significantly affected by storage time.

**NMR METABOLOMICS ANALYSIS**  
Samples (H2–H6: T0, T30, T120) were randomized before sample preparation and analysis. We diluted 50 μL CSF in 200 μL heavy water (D2O) (99.999% D). We added 25 μL of 8.8 mmol/L TSP-d4 [3-(trimethylsilyl)propionic acid-d4 sodium salt, 99% D] stock solution in D2O to 250 μL CSF to a final concentration of 0.8 mmol/L TSP as internal standard and as chemical shift reference (80.00). The TSP-d4 stock solution was prepared from dry TSP-d4. The pH was adjusted (7.0–7.1) by adding phosphate buffer (9.7 μL of 1 mol/L stock solution) to a final concentration of 35 mmol/L (27). Finally the sample (284.7 μL) was trans-
ferred to a Shigemi microcell NMR tube for measurements. Each sample was analyzed once.

One-dimensional 1H NMR spectra were acquired on an 800-MHz Inova (Varian) system equipped with a 5-mm triple-resonance, Z-gradient HCN cold probe. Suppression of water was achieved using Watergate (delay, 85 s)\(^{28}\). For each spectrum, 256 scans of 18 000 data points were accumulated with a spectral width of 9000 Hz. The acquisition time for each scan was 2 s. An 8-s relaxation delay was used between scans. Before spectral analysis, all acquired free induction decays were zero-filled to 32 000 data points, multiplied with a 0.3-Hz line-broadening function, Fourier-transformed, and manually phased. Calibration of the chemical shift scale was done on the external reference standard TSP-d\(_4\) by use of ACD/SpecManager software (Advanced Chemistry Development). Spectra were transformed to MatLab, version 7.6 (R2008b) (Mathworks), for further analysis.

NMR spectral data were preprocessed by baseline correction using the asymmetric least squares method \(^{(29)}\) and aligned with the correlation optimized warping method \(^{(30)}\). Each spectrum was divided (along the chemical shift axis) into equally sized bins (0.04 ppm), and each data point was averaged over each bin. The areas of the bins were summed to provide an integral so that the intensities of the peaks in such defined spectral regions could be extracted. Each NMR spectrum was reduced to 210 variables, calculated by integrating regions of equal width (0.04 ppm) corresponding to the regions of 0.7–9. To remove effects of variation in water resonance suppression, spectral regions between 4.4 and 5.4 were removed. All spectra thus reduced were normalized to unit area.

The data were further processed by supervised vast scaling, to determine group-specific scaling factors \(^{(31)}\). To visualize possible systematic variation, grouping, trends, and outliers, we applied PCA to the entire data set. To remove biological (patient-to-patient) variation, data were further mean-centered per patient and vast-scaled.

**LC-MS/MS AMINO ACID ANALYSIS**

CSF samples (H2–H6: T0, T30, T120) were prepared in triplicate as described \(^{(21)}\). We injected 1 \(\mu\)L of each reaction mixture in duplicate on an Acquity UPLC\(^\text{TM}\)
Stability of CSF Proteome and Metabolome

Results

PROTEOMICS ANALYSIS

The Orbitrap-MS/MS shotgun analysis resulted in a list of 55,421 peaks, from which 5780 peptides were identified. All identified peptides from the Orbitrap-MS/MS data and the 10,000 most intense QTOF-MS peaks (complete peak matrix) were used for unsupervised multivariate statistical analysis (PCA). No trend with respect to delay before storage was visible (Figs. 1A and 2A). Run-to-run variability on the LC-MS/MS platforms for proteomics was within 20% with respect to peak area and within 0.5% with respect to retention time (21). PCA shows that biological variation is more prominent than the effect of time at room temperature between sampling and freezing, since data points clustered according to the individual patients rather than according to time points (Figs. 1B–1D and 2B–2D).

ANOVA comparison by time variance on the Orbitrap-MS/MS data resulted in 56 peaks (6 identified peptides) with a $P$ value $< 0.01$, which is well below 554, the number of peaks that would receive this $P$ value when comparing identical samples (null hypothesis) due to chance alone (1% of all 55,421 detected peaks). NSC analysis pointed also to only random differences between the time groups, with a cross-validation error of 0.5. This led to the conclusion that there was no significant discrimination between the samples stored at $-80$ °C immediately after centrifugation and samples left at room temperature for 30 or 120 min before being frozen and stored for the observed proteins.
NSC analysis of the QTOF-MS data revealed that differences between T0 and T120 were random, with a double cross-validation error of 0.5. Comparison of T0 vs T30 by NSC reached a minimal average cross-validation error of 0.34 and resulted in only 11 selected peaks. PCA on the NSC-selected peaks (T0 vs T30) from the QTOF data showed no clear discrimination but a weak tendency of clustering according to time groups (Fig. 3A).

The concentrations of 2 proteins in CSF, albumin and cystatin C, were measured by targeted mass spectrometric analysis in the SRM mode. These proteins are exemplary of the vast majority of CSF proteins, which we found remained unchanged after 120 min at room temperature. Albumin was chosen as it represents the largest part of CSF total protein, a parameter that is often used in CSF-based diagnosis of disease, and cystatin C was chosen as a protein that is sensitive to storage conditions (1, 2). Concentrations of cystatin C and albumin were calculated based on the measured ratios of the corresponding spiked isotope-labeled internal peptide standards to their biological counterparts, confirming that variation between the different time points was not statistically significant (online Supplementary Tables S3 and S4). The measured concentrations of both proteins were both found to agree with reported CSF concentrations (32, 33). The albumin concentrations measured by SRM were also in agreement with albumin concentrations measured by standard clinical chemistry techniques (Table 1). Additionally, we assessed the trypsin cleavage efficiency by monitoring the release of a tag from the lysine end of the cystatin C peptide during the regular digestion procedure. After the overnight digestion procedure, the entire peak in the LC-MS data corresponding to the peptide including the tag had completely disappeared.
indicating that a complete digestion had taken place. The observed relative SDs for the cystatin C measurements were <10%, and those for albumin were <4%. Technical variability with sample pretreatment was <4%, and without sample pretreatment, <2% (Table 2).

### Table 2. Relative standard deviations (RSDs) of the SRM measurements of cystatin C and albumin.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Samples, n</th>
<th>RSD of cystatin C, %</th>
<th>RSD of albumin, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>3</td>
<td>5.20</td>
<td>2.19</td>
</tr>
<tr>
<td>H2</td>
<td>3</td>
<td>6.62</td>
<td>1.78</td>
</tr>
<tr>
<td>H3</td>
<td>3</td>
<td>7.72</td>
<td>3.13</td>
</tr>
<tr>
<td>H4</td>
<td>3</td>
<td>7.55</td>
<td>3.05</td>
</tr>
<tr>
<td>H5</td>
<td>3</td>
<td>9.15</td>
<td>2.59</td>
</tr>
<tr>
<td>H6</td>
<td>3</td>
<td>2.72</td>
<td>1.51</td>
</tr>
<tr>
<td>Technical variation, sample pretreatment</td>
<td>3</td>
<td>3.74</td>
<td>2.34</td>
</tr>
<tr>
<td>Technical variation, mass spectrometry measurement</td>
<td>4</td>
<td>0.82</td>
<td>1.74</td>
</tr>
</tbody>
</table>

* Delayed storage at −80 °C caused slightly higher RSDs of protein concentrations in individual patients than did technical controls of measuring a single sample multiple times. The values also depend on the measured protein. Average RSD for cystatin C in the 6 patient samples within 2SD = 6.25 ± 4.99; average RSD for albumin in the 6 patient samples within 2SD = 2.40 ± 0.66.

**Discussions**

We present a study of the stability of the measured proteome and metabolome of human CSF samples are left at room temperature for up to 2 h between lumbar puncture and storage at −80 °C, to mimic delayed storage in clinical routine practice.

Unsupervised multivariate statistical analysis (PCA) showed that patient-to-patient variation was the most prominent, overriding variation that was due to delayed storage time. After variable selection based on preclassification of the samples according to delay time, we found that only 2 peptides and 2 metabolites changed significantly over time among approximately 6000 detected peptides (2755 unique peptides) and 88 detected metabolites. Our results demonstrate that human CSF prepared according to the described procedure is suitable for proteomics and metabolomics analysis even when left at room temperature for 2 h, provided that all cells have been removed by centrifugation. Quantification of albumin and cystatin C by targeted mass spectrometry in the SRM mode using stable isotope–labeled internal standard peptides showed that there is no statistically significant difference over 2 h of delay time, confirming our proteomics results.

Another study on the stability of the proteome in CSF at room temperature pointed in the same direction, with the detection of only 2 polyproteptides that changed concentrations after storage (32). These samples were contaminated with blood, however, since both polyproteptides were derived from hemoglobin. Another study showed that blood contamination decreases the stability of the CSF proteome (11), corroborating our earlier results (21). One explanation for the decreased level of the 2 unidentified peaks in the proteome analysis is the possible adsorption to the vial surface, e.g., via hydrophobic or van der Waals interactions (21, 33, 34). Metabolomics revealed increased concentrations of threonic acid after storage at room temperature. This increase may be caused by oxidative degradation of ascorbic acid (35–37), because here the ascorbic acid levels were slightly decreased with increased time at room temperature. This decrease was too small to be a significant factor, however. Interest-
ingly, threonic acid decreased in CSF containing white blood cells (21), which might be due to further metabolism of the acid by enzymes released from white blood cells. The concentration of threonic acid was too low for NMR detection.

The stability of metabolites and proteins measured with the most common analytical profiling methods is an important factor to take into consideration when handling biofluids and designing biomarker studies. A previous study from our team shows that the biological variation of some proteins and peptides has large variability (sometimes exceeding 100%), which requires high discriminating power for compounds considered as biomarker candidates (38).

In conclusion, we assessed the stability of CSF using 5 different analytical platforms, showing that overall there are very minor changes in either peptides (2 of approximately 6000 in trypsin-digested CSF) or metabolites (1 metabolite of 93). Earlier studies showed that blood or white blood cell contamination reduces CSF stability considerably, emphasizing the importance of the initial centrifugation step. Because we did not add antioxidants, we cannot draw conclusions about other oxygen-sensitive metabolites such as the catecholamines. The observed increase in threonic acid over time, however, indicates that oxygen-sensitive metabolites require specific protective measures during sample preparation and storage.

**Author Contributions:** All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.

**Authors’ Disclosures or Potential Conflicts of Interest:** Upon manuscript submission, all authors completed the Disclosures of Potential Conflict of Interest form. Potential conflicts of interest:

- **Employment or Leadership:** None declared.
- **Consultant or Advisory Role:** None declared.
- **Stock Ownership:** None declared.
- **Honoraria:** None declared.

**Research Funding:** The study was performed within the framework of the Top Institute Pharma project number D4-102. T. Rosenling, the project BioRange 2.2.3 from the Netherlands Proteomics and the Netherlands Bioinformatics Center; C. Christin, the project...
BioRange 2.2.3 from the Netherlands Proteomics and the Netherlands Bioinformatics Center; P. Horvatovich, the project BioRange 2.2.3 from the Netherlands Proteomics and the Netherlands Bioinformatics Center; R. Bischoff, the project BioRange 2.2.3 from the Netherlands Proteomics and the Netherlands Bioinformatics Center.

Expert Testimony: None declared.

Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, or preparation or approval of manuscript.

Acknowledgments: The authors thank Prof. Dr. Rogier Hintzen from the Department of Neurology, Erasmus University Medical Center, Rotterdam, The Netherlands, for providing the CSF samples.

References