Master’s Thesis of Engineering

Web Portal for the assay analytical validation of biomarkers abided by multinational integrative guidelines

다기관 가이드라인 준수 바이오마커 분석 검증 웹포탈

February 2018

Graduate School of Engineering
Seoul National University
Bioengineering Major

Jae Nyeon Kim

Web Portal for the assay analytical validation of biomarkers abided by multinational integrative guidelines

Youngsoo Kim

Submitting a master's thesis of Public Administration

February 2018

Graduate School of Engineering
Seoul National University
Bioengineering Major

Jae Nyeon Kim

Confirming the master's thesis written by
Jae Nyeon Kim
February 2018

Chair _____(Seal)
Vice Chair _____(Seal)
Examiner _____(Seal)

Web Portal for the assay analytical validation of biomarkers abided by multinational integrative guidelines

지도교수 김 영 수

이 논문을 공학석사 학위논문으로 제출함
2018년 2월

서울대학교 대학원
공과대학 (협)바이오엔지니어링 전공

김 재 년

김 재 년 의 공학석사 학위논문을 인준함
2018년 2월

위 원 장 _____(인)
부위원장 _____(인)
위 원 _____(인)

Abstract

Multiple reaction monitoring mass spectrometry (MRM-MS) involves monitoring a multiplexed assay of peptides and associated transitions in mass spectrometry runs. Implemented as a freely available, open-source tool in the platform independent Java programing language, assay Portal computes analytical measures as recommended integrated multinational guidelines from FDA,EMA,KFDA which applies to clinical assays. Computed measures include; calibration curve, specificity, selectivity, interference, sensitivity, carryover, precision, accuracy, quality control samples, matrix effect, recovery, dilution integrity, and stability. Assay Portal streamlines assay development analytical workflow and therefore minimizes error predisposition. Assay Portal may also be used for performance estimation for MRM-MS assays. Assay Portal is available from http://pnbvalid.snu.ac.kr with sign up procedure only available by manual confirmation.

Keyword : Method Validation, Portal, Database, Biomarker, Proteomics, MRM
Student Number : 2016-21166

Table of Contents

Chapter 1. Introduction

Clinically validated biomarkers require an analytical validation step of biomarker assay in order to reach the clinical applications (ref). Analytical validation involves confirming that the method used for the biomarker measurement is accurate, precise, specific, robust, and stable over time [1–5]. In this step, all biomarkers of several analytical validation results should be reported in a detailed and transparent manner.

Over the years, there has become an increasing need for highly multiplexed protein panels to help expedite the validation of putative protein biomarkers and to help improve the diagnostic/prognostic accuracy of disease assessment [6,7]. This requires an accurate quantification of multiple protein biomarkers at once, increasing the use of multiple reaction monitoring–mass spectrometry (MRM–MS) assays for clinical applications.

However, it is difficult to manually interpret and evaluate several analytical validation procedures because the ㅜMRM–MS assay produces simultaneous measurements of thousands of transitions (i.e., light, native, or endogenous; heavy, stable isotope–labeled standard, or internal standard) corresponding to the quantitative values of multiple protein biomarkers.

Currently, the MRM–MS data analysis can be partially accomplished with vendor–dependent software (MassHunter

Quantitative Analysis, Agilent; MultiQuant, ABSciex; Pinpoint, Themo Scientific) or with vendor-independent programs, such as Skyline (ref). Overall, these software packages are generally dedicated to a preliminary analysis of the mass spectrometric spectral data and the transitions and enable the user to verify and edit the peak selection/integration. However, none of the available software has the function to quickly evaluate whether thousands of transitions analyzed by MRM-MS assay have been analytically validated or not.

To address this unmet need, a web Portal that automatically evaluates the analytical validation of multiple protein biomarkers with MRM-MS assay. The analytical validation items configured in the Portal are designed to meet the requirements of 3 sets of guidelines [US Food and Drug Administration, (FDA), European Medicines Agency (EMA) and Korea FDA (KFDA)]. These items covered such aspects as calibration curve, specificity, sensitivity, carryover, precision, accuracy, matrix effect, recovery, dilution integrity, stability, and quality control (QC) of samples and frequency.

Access to Portal content with Skyline-derived MRM-MS data allows users to select each country-specific entry through customizations designed to adhere with the analytical validation according to country-specific guidelines, and automatically evaluate whether the data meets the validation practices and performance specifications. This web Portal also displays

analytical validation results in a table form.

The PORTAL, web Portal centralizes all calculations for analytical validation into a single tool, providing a significant reduction in time, effort, and errors that can occur through manual processing. This ultimately facilitate an attempt to apply the MRM-MS assay to clinical implementations by easily access to the analytical validation process of multiple protein biomarkers.

Chapter 2. Method & Material

2.1. Skyline format file

Format of skyline-exported files are critical for accurate validation. In normal condition, technician uses Skyline to checkout area of all detected material to ensure results data contains precise values. Technician must modify data column order and name of analyte prior to exporting results data as csv files. Throughout all the experiment categories, “Protein Name”, “Peptide Sequence”, “Replicate Name”, “Replicates”, “Precursor Charge”, “Product Charge”, “Fragment Ion”, “light Precursor Mz”, “light Product Mz”, “light Retention Time”, “light Area”, “heavy Precursor Mz”, “heavy Product Mz”, “heavy Retention Time”, “heavy Area” are required in the exported csv file in exact column order.

Name of Analyte varies to each categories, which the contents in the name value generally follows order and separation rules if needed: “Temperature value-day value-Matrix value_concentration point value”. “_” is used to separate and indicate matrix number to concentration point and “- “ is used to separate date and temperature value to matrix value. These naming rules are crucial, as naming values are the only key to identify data from each other and 12 validation categories requires different name values for precise

calculation and concentration points or number of matrix samples are only available by analyte name. Portal offers an example template on PORTAL for user to follow.

2.2. Database Structure

Database is constructed based on MySQL Server 5.7 Community version. It contains four tables: User, information, experiment and experiment data table. User table contains ID, password and other minimum personal information. Information table contains experiment information which user has to insert like sample type, title, target instrument, target source, sample description, organism and experiment date, and automatically filled user ID and experiment ID to manage experiments from each other. Experiment table documents every attempt of file uploads to PORTAL. It stores user ID, experiment ID, category of uploaded file and upload time. It also contains validation results for each administration, which changes with results of uploaded file. As PORTAL only calculates latest uploaded file, so for each category under same experiment ID for a user, the latest stored in experiment table will be called and used. Experiment data table stores all the data from uploaded file. Each entry is linked to exact match to experiment table data for later calculation.

Figure 1. Database Schema

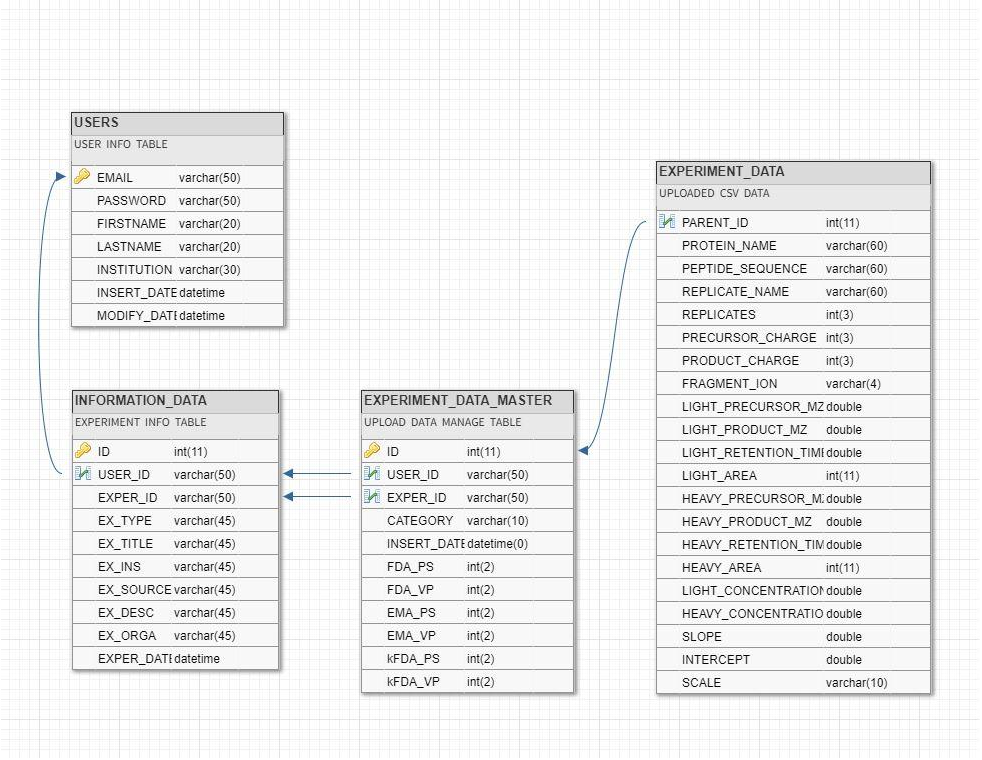


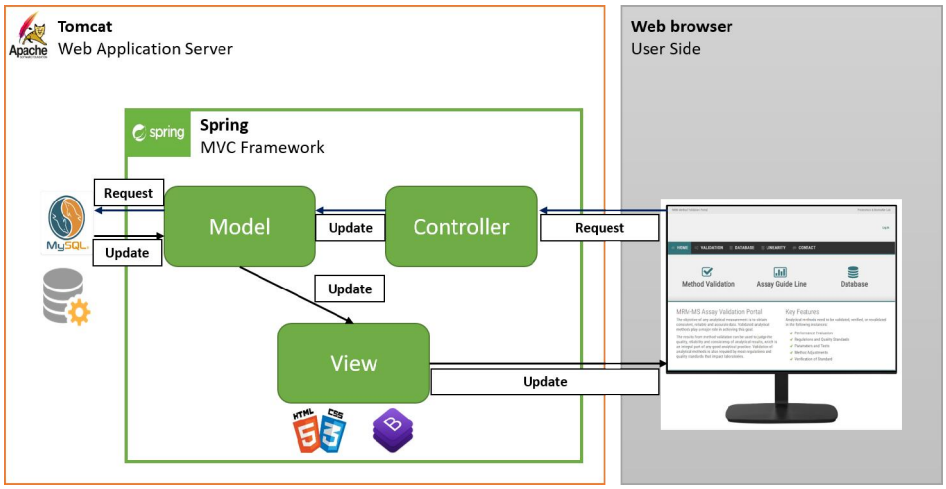
Figure 2 Query as XML by myBatis

```
<select id="selectDcList" parameterType="String" resultType="String">
  SELECT distinct RIGHT(REPLICATE_NAME,2) as dcName
  FROM experiment_data
  WHERE PARENT_ID=#{parentId}
  AND RIGHT(REPLICATE_NAME,2) REGEXP '^[0-9]+$'
</select>

<select id="selectExperList" parameterType="String" resultType="String">
  SELECT EXPER_ID
  FROM INFORMATION_DATA
  WHERE USER_ID=#{USER_ID}
</select>

<select id="selectExperData" parameterType="String" resultType="java.util.HashMap">
  SELECT PROTEIN_NAME,PEPTIDE_SEQUENCE,FRAGMENT_ION,PRODUCT_CHARGE,REPLICATE_NAME,ROUND(
    FROM experiment_data
   WHERE PARENT_ID= #{PARENT_ID}
  GROUP BY PROTEIN_NAME,PEPTIDE_SEQUENCE,FRAGMENT_ION,PRODUCT_CHARGE,REPLICATE_NAME;
</select>

<select id="selectExperData_1" parameterType="String" resultType="java.util.HashMap">
  SELECT PROTEIN_NAME,PEPTIDE_SEQUENCE,FRAGMENT_ION,PRODUCT_CHARGE,REPLICATE_NAME,ROUND(
    FROM experiment_data
   WHERE PARENT_ID= #{PARENT_ID}
  GROUP BY PROTEIN_NAME,PEPTIDE_SEQUENCE,FRAGMENT_ION,PRODUCT_CHARGE,REPLICATE_NAME;
</select>
```

2.3. Server Design

System uses Spring open source framework to buildup the server. Java controller to handle request and mapping, JavaServer Pages for dynamic web page with ajax for asynchronous web application. Server implanted Mybatis framework to handle SQL statements for calculation and storing data. Gradle build automation system is used to for WAR build and project configuration. Bootstrap framework is used for web application UI design.

Figure 3. Server Configuration



2.4. Calculation & Validation

To insure the calculation constancy, Portal performed data calculation and validation using only SQL query. Calculation method and validation criteria are performed separately as following. Heavy Area and Light Area is used to calculate Peak area ratio(PAR) and average, sample standard deviation, coefficient of variance of PAR are calculated among replicates of each calibrator and matrix.

2.4.1 Calibration Curve and linearity

Linearity calculation is accessed by using uploaded calibration data and manual input of expected concentration of each calibrator and its middle point(IS) concentration. Concentration ratio, which is calibrator concentration divided by middle point concentration, is used as X factor while PAR is used as Y factor of the linearity. Slope, intercept and Coefficient of determination is calculated by simple linear regression. For each target, X, Y factor can be transformed in to normal, Log2 or Log10 scale as user demands and its linear range also can be manually selected by user for final linear function. Slope, intercept and scale of each linear range included in the uploaded raw data is stored in the database as default which alters by every attempts user changes its linearity range and scale. Final selection of range and scale is used for the later calculation of

concentration of each targets calibrator. Concentration calculation is done by input PAR to average linear function of each target. Concentration ratio will be acquired with the process and multiply the value with expected IS concentration will generate measured concentration. Bias is calculated by comparing expected concentration with measured concentration which accessed by input PAR to linear function of each targets matrix. For all administrations standard of performance specification(PS) validation, bias is allowed within 15% for all calibrators while within 20% at Lower Limit of Concentration(LLOQ), lowest concentration on the reverse calibration curve, for above 75% of calibration standards. EMA specifies 20% bias at Upper Limit of Concentration(ULOQ), highest concentration on the reverse calibration curve, as well, while kFDA standard for amount of bias is above 50% of calibration standards. For validation practices(VP), all administration specifies method needs to include blank, zero and minimum of 6 points of calibration standards for every matrix and at least 5 different matrix for every target.

2.4.2 Sensitivity

Sensitivity data is assessed by calculate signal to noise(S/N) ratio of LLOQs and zero sample PAR of each target matrix. Accuracy is also calculated comparing measured concentration

and expected concentration of LLOQ. PS standard indicates that S/N has to be at least 5 and accuracy between 80% to 120%. At least 5 sample matrices should be measured for VP.

2.4.3 Carryover

Carryover of internal standard(IS) and LLOQ should be calculated to validation carryover. For each replicates of target matrix, concentration of zero sample and blank sample after ULOQ, ULOQ needs to be calculated with linear function. Zero sample carryover requires dividing zero sample concentration by expected LLOQ concentration, while blank sample carryover requires dividing blank sample PAR to ULOQ PAR, then calculate average, standard deviation, coefficient of variation among target matrix. Only EMA and kFDA have validation standards, average carryover should be under 20% in zero sample and 5% in blank sample for PS, existence of blank sample after ULOQ for VP.

2.4.4 QC

Skyline raw data regarding QC and expected concentration of QC sample needs to be manually input via webpage. In the guideline, QC concentrations are suggested to distributed at least around LLOQ, middle range and ULOQ, so I suggest to set QC low concentration at LLOQ, midrange at (LLOQ+ULOQ)/2

and high concentration at 0.9*ULOQ. Calculating accuracy by dividing measured concentration of each QC sample to expected concentration for each replicates of target matrix. FDA and EMA PS guideline indicates that accuracy of at least 4 of every QC sample needs to within 85% to 115% and not 2 or more sample beyond 15% in the same concentration. VP guideline shows there must be at least 3 concentration at low middle, high concentration. It also indicates that number of QC sample analyzed during a batch should represent more than 5% of the total number of patient samples, which is a criteria cannot validate with data uploaded. Webpage for validation results will indicate user to check this criteria manually.

2.4.5 Precision and Accuracy

These two categories are assessed by analyzing QC samples. VP standard of 3 administrations indicates that At least 5 replicates and in 5 days analyzation of more than 3 concentration of QC sample data is required to pass the validation. Within run precision and accuracy is calculated by averaging concentration of replicates of each target QC concentration on each day while between run is calculated by averaging first first run of each target QC concentration of all days. Average, sample standard deviation, CV of results and accuracy assessed by dividing measured concentration by

expected concentration. Both CV and accuracy values needs to be within 15% while within 20% on LLOQ.

2.4.6 Recovery

Recovery is assessed by comparing PAR of target matrices analyte added/extracted and true concentration of the analyte in solvent on each QC concentrations. There is no exact guideline for recovery category by 3 administrations but to control recovery of at least 3 concentrations with consistent and reproducibility, System set a limit to the recovery result to stay within 20% of nominal value according to accuracy category.

2.4.7 Matrix Effects

For matrix effects validation, EMA and kFDA requires definite standard than FDA, which only state that the corresponding experiment should be done with calibration curve spiked into matrix vs neat solution. EMA and kFDA both requires at least 6 different matrices samples spiked with analyte post-extraction to analyte spiked into neat solution tested at 3 times of LLOQ concentration and high concentration near ULOQ. CV of results should stay within 15% to pass performance specification. Dividing PAR of spiked samples to neat solution calculates matrix effect and averaging all matrix results in same concentration provides final values for validation.

2.4.8 Specificity

Guideline for specificity is similar to recovery guideline by three administrations. At least 5 different matrix samples needs to be tested, response from a potential interference should be under 20% at LLOQ concentration for the analyte while under 5% in blank sample to pass all three administrations guideline. Portal assessed validation values by calculating concentration of zero, blank and LLOQ sample. Dividing concentration in zero samples by measured LLOQ concentration provides specificity for zero sample, while dividing PAR of blank sample to LLOQ concentration provides specificity for blank sample.

2.4.9 Dilution Integrity

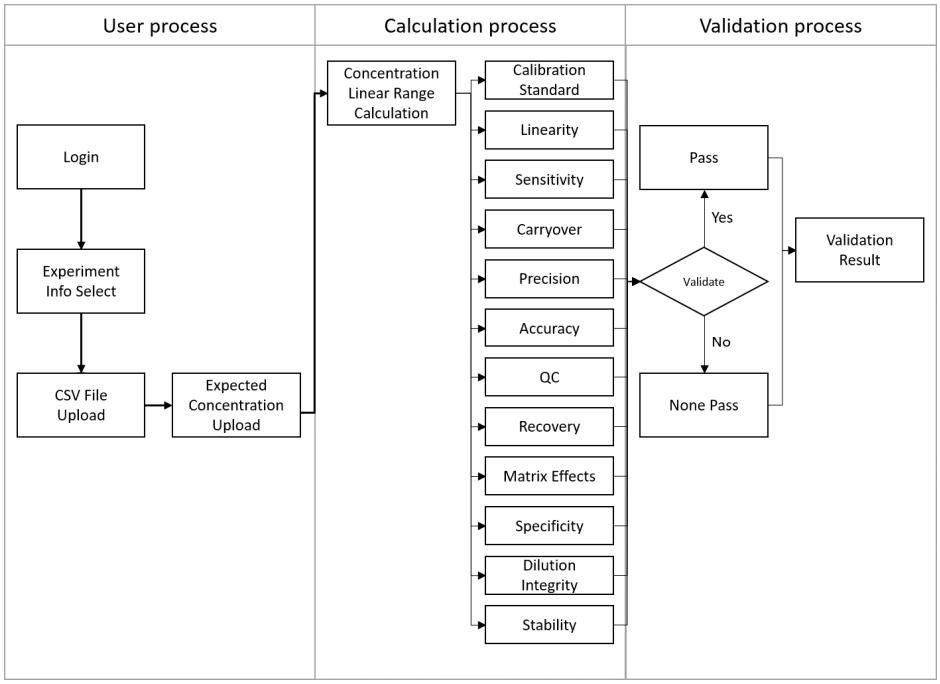
Prior to validate dilution integrity, user has to input an initial concentration of experiment, which needs to be higher than ULOQ concentration. Diluting to at least 5 concentrations is needed to pass the validation standard of guidelines. Analyte name set for dilution integrity data contains dilution factor, which will multiply with measured concentration for each diluted sample called dilution corrected concentration. Dividing corrected concentration minus neat concentration to neat concentration provides concentration change ratio effected by dilution, which the change ratio of results should not over 15%

to pass the guideline.

2.4.10 Stability

Standard for stability validation is an vague standard which cannot be define as a certain value. Validation practice standard of all administrations recommends experiments should be done with freeze and thaw, short-term, long term and sample stability check. Short, long term is a vague standard that not an exact value is offered for validation. For this part user must develop experiment method with own vision. At least 3 aliquots of low to high QC concentration must be measured to pass the guideline. For performance specification, mean concentration at each level must varies between 15% of nominal concentration, which is assessed by comparing measured concentration results to initial day of experiment result, defined as 0 day result. Server will identify 0 day results by analyte name, and compare the results any other analyte of corresponding experiment data on every QC concentration.

Figure 4.Overall Process Diagram



Chapter3. Result

PORTAL is an optimized website to validate technicians skyline generated data file. Whole procedure is straightforward and intuitive.

Login is required to validate experiment data on PORTAL and sign up for PORTAL is current only available after administrator authorization due to limit access to the website.

After login, user will be available to either insert or select pre-inserted experimental information associate with following csv data. Experimental Information stores sample type, title, target instrument, target source, sample description, organism and experiment date to distinguish each experiments. An Experiment info ID will be generated using date time of information entry, experiment date and sample type. Selecting desired experiment ID will lead to the upload procedure. When developing PORTAL, its fundamental design were based on validation standards by three administrations, hence uploading categories for Calibration Curve(Forward and Reverse), Specificity, Sensitivity, Carryover, Precision/Accuracy, QC, Matrix Effect, Recovery, Dilution Integrity and Stability is provided and uploaded data are stored in the database. User can upload one to multiple file at the same time, while the system calculates and validates latest uploaded data for each category. It is critical that upload files are in csv format, and column

names, contents and its order have to be in exact regulation or the calculation will not be correctly done. To avoid error input, an example csv format is available for download.

On normal occasion, before validate other categories, technician precedes reverse calibration curves related experiments. For every batch of targeted material, multiple points of concentration range including standard sample, buffer blank and curve blank needs to be measured. With internal standard(IS) injected in the samples, light and heavy area associated to IS and standards depending on forward/reverse curve are generated by skyline and export as csv file. Database will analyze how many concentration points are in the csv file, and following the upload procedure will ask user to insert theoretical concentration of each point. With theoretical concentration updated, system will calculate slope, intercept and r squared value and displayed in table form. User can also chose Normal, Log2 or Log10 scale calculation for linear regression data and the latest choice of scale will affect all the calculation beyond. QC sample and dilution integrity also needs insertion of its theoretical concentration. Figure of linear regression is also available for user to check if there is any irregular concentration point in the concentration range despite R squared value.

Depending on each categories, peak area ratio(PAR) or concentration values are calculated and displayed as table form

on each webpage. For most categories, average, sample standard deviation and coefficient of variation value of PAR or concentration are calculated as those are the values to validate method. While pages are called from server and displayed, database also validates if the current categories passes the validation standard of three administration and store it along with uploaded data. After examine all the categories, user can check if the categories passes experiment and result criteria by displaying Pass, Not Pass or Not Addressed if the administration did not specify a certain standard.

To test the performance and data integrity of the PORTAL, uploaded skyline generated raw file of our previous experiment following three administrations standard and compared PORTAL generated results to manual calculated results. Deviation is observed after three decimal point, which leaded by decimal point difference between PORTAL and manual during calculation. Throughout the result, observed deviation was minimal that did not affect the validation result of our previous work.

Figure4. Main page of Validation portal

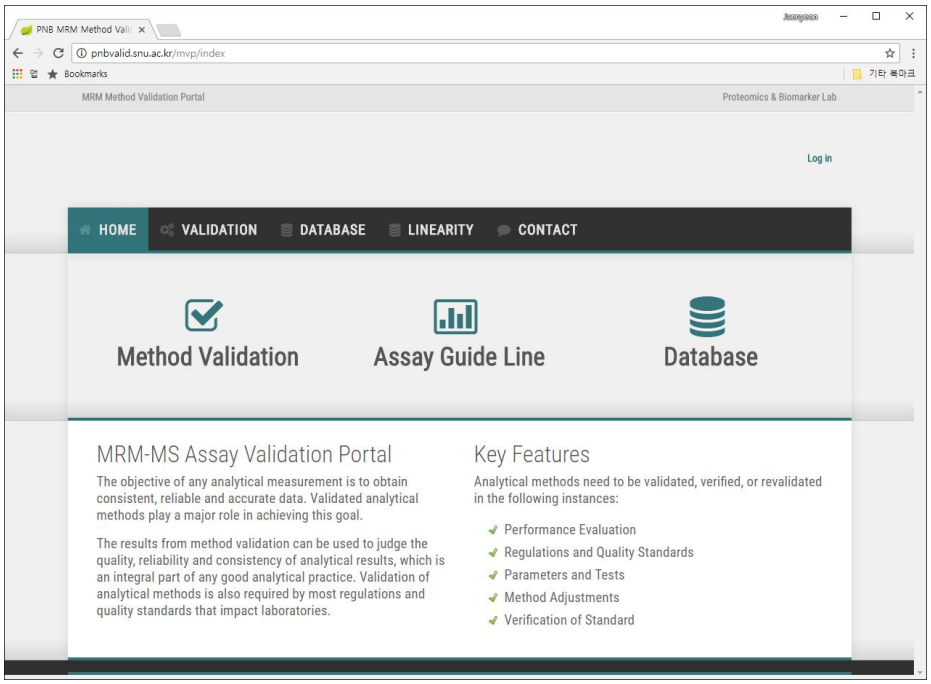


Figure 5. Instruction and Sample data Page

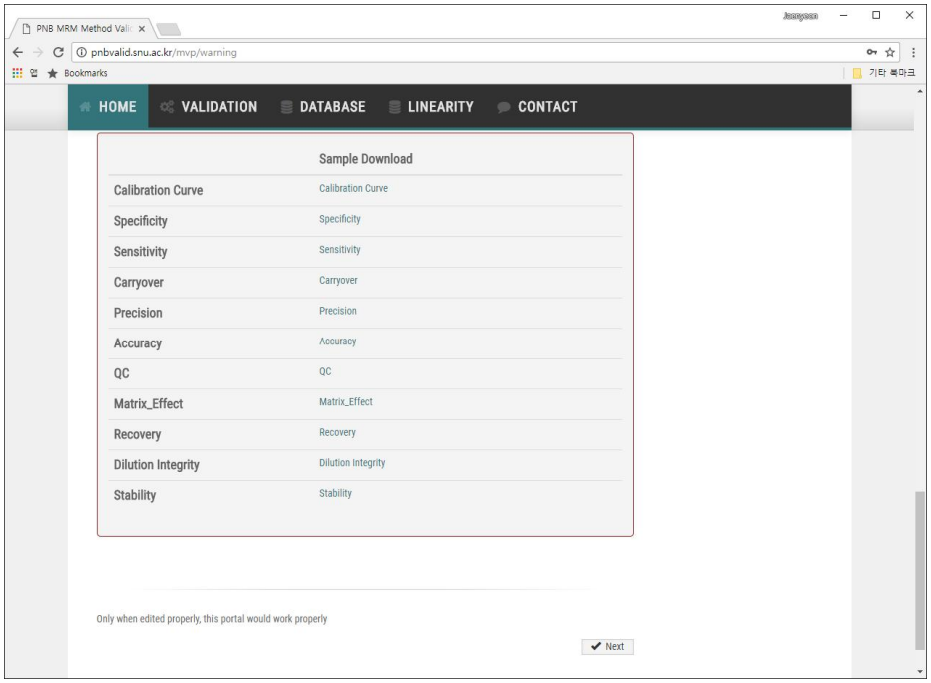


Figure 6. Experiment information page

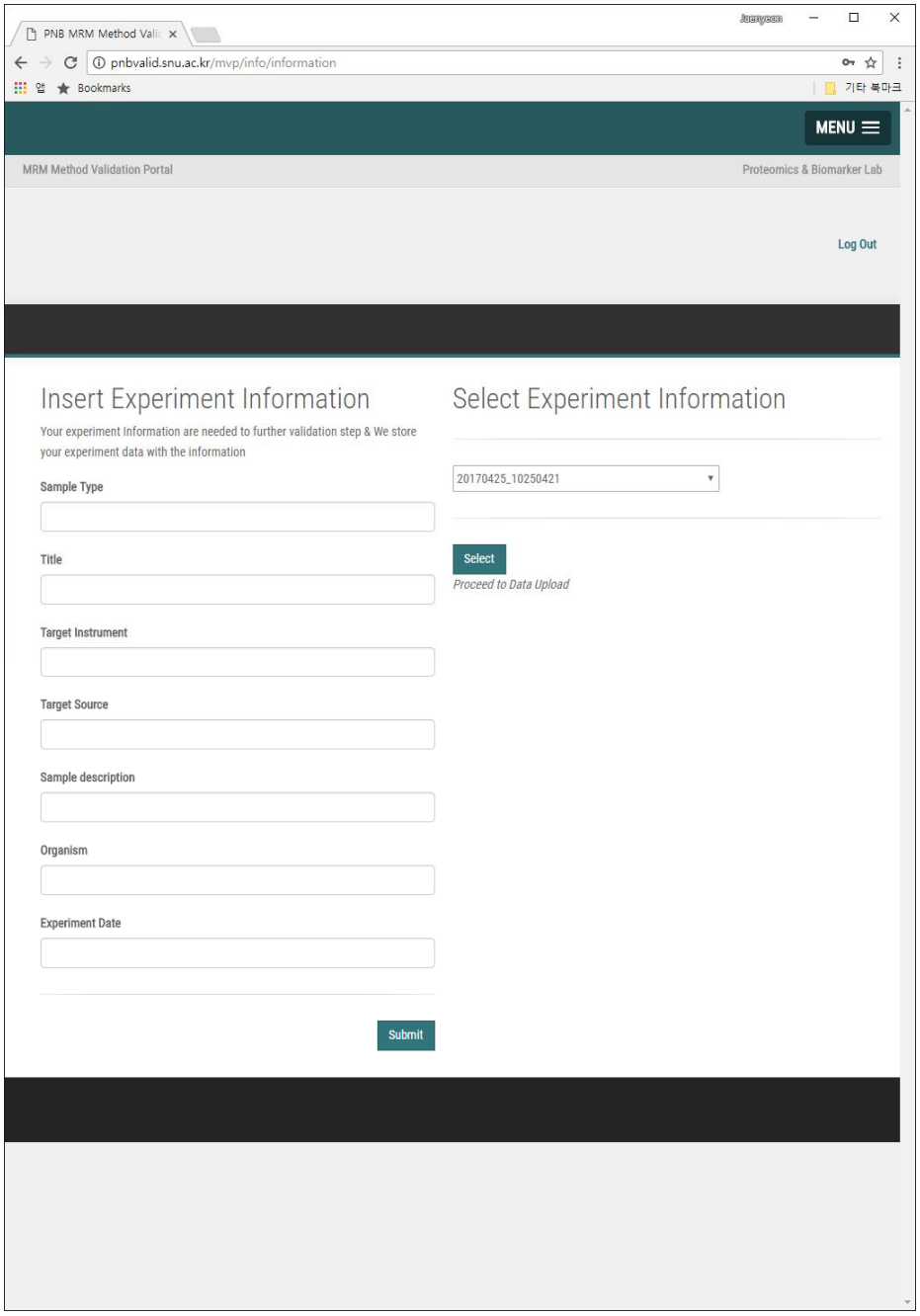


Figure 7. Data Upload Page

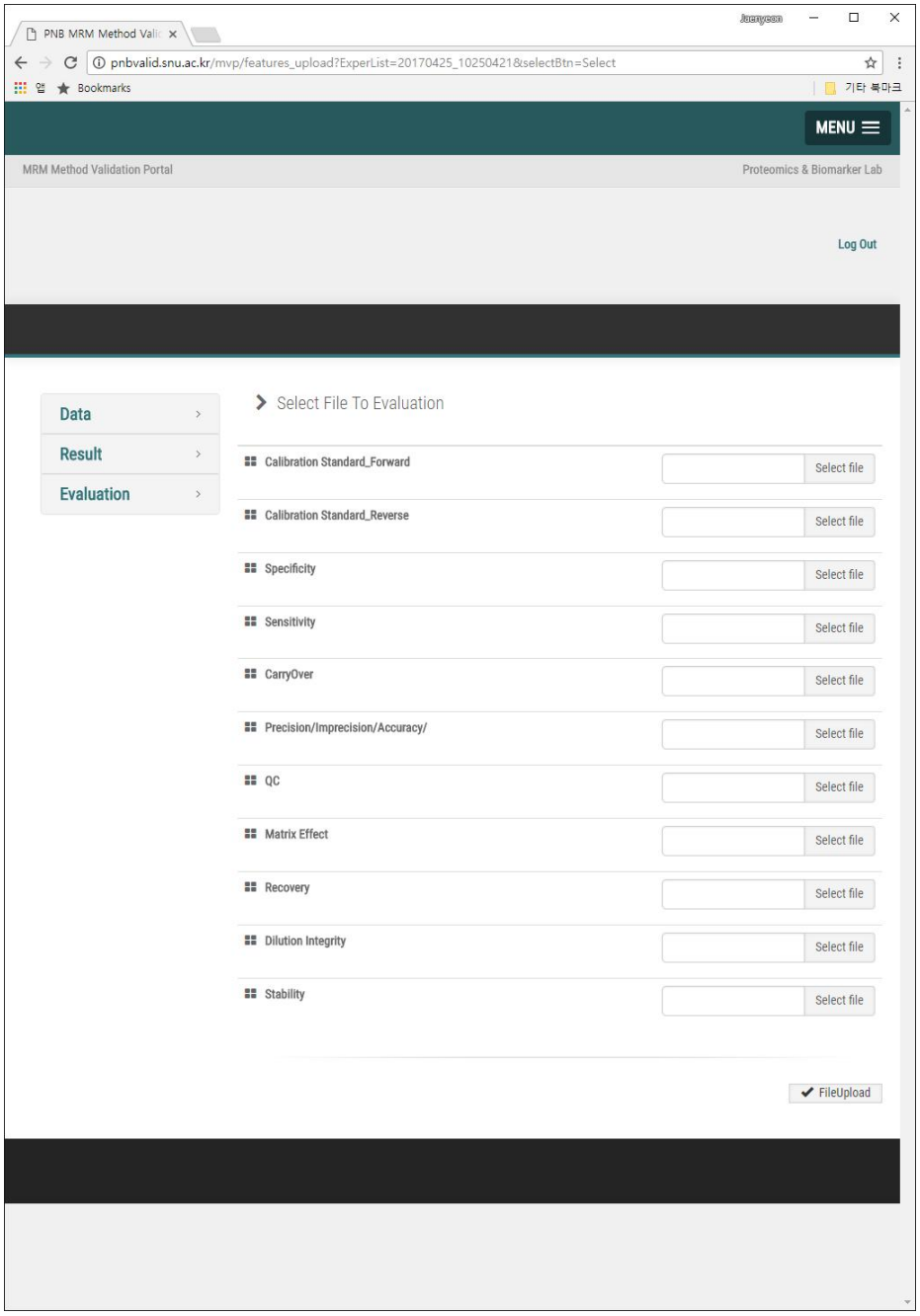


Figure 8. Dilution Concentration Page

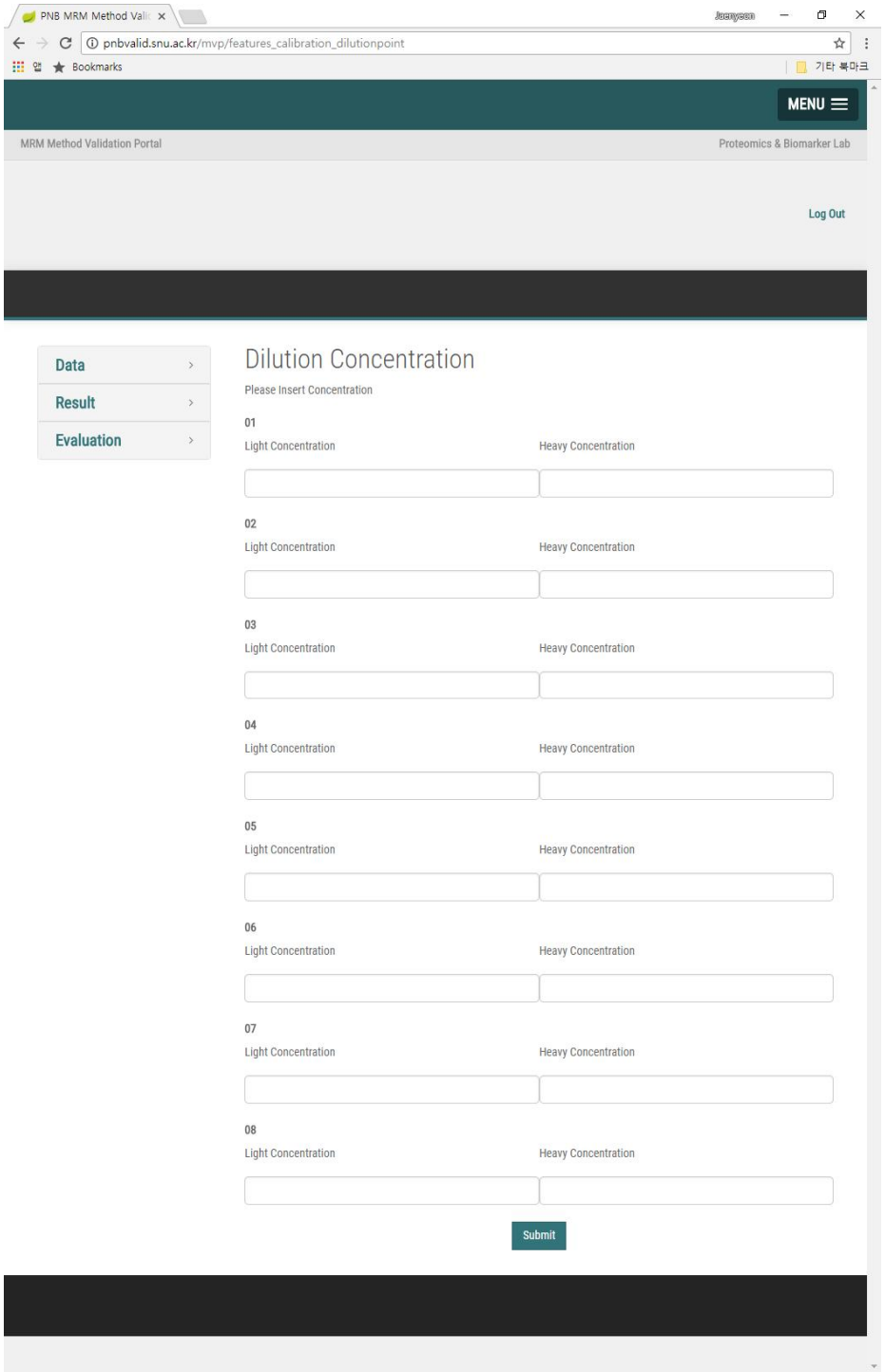


Figure 9. Linearity Selection Page

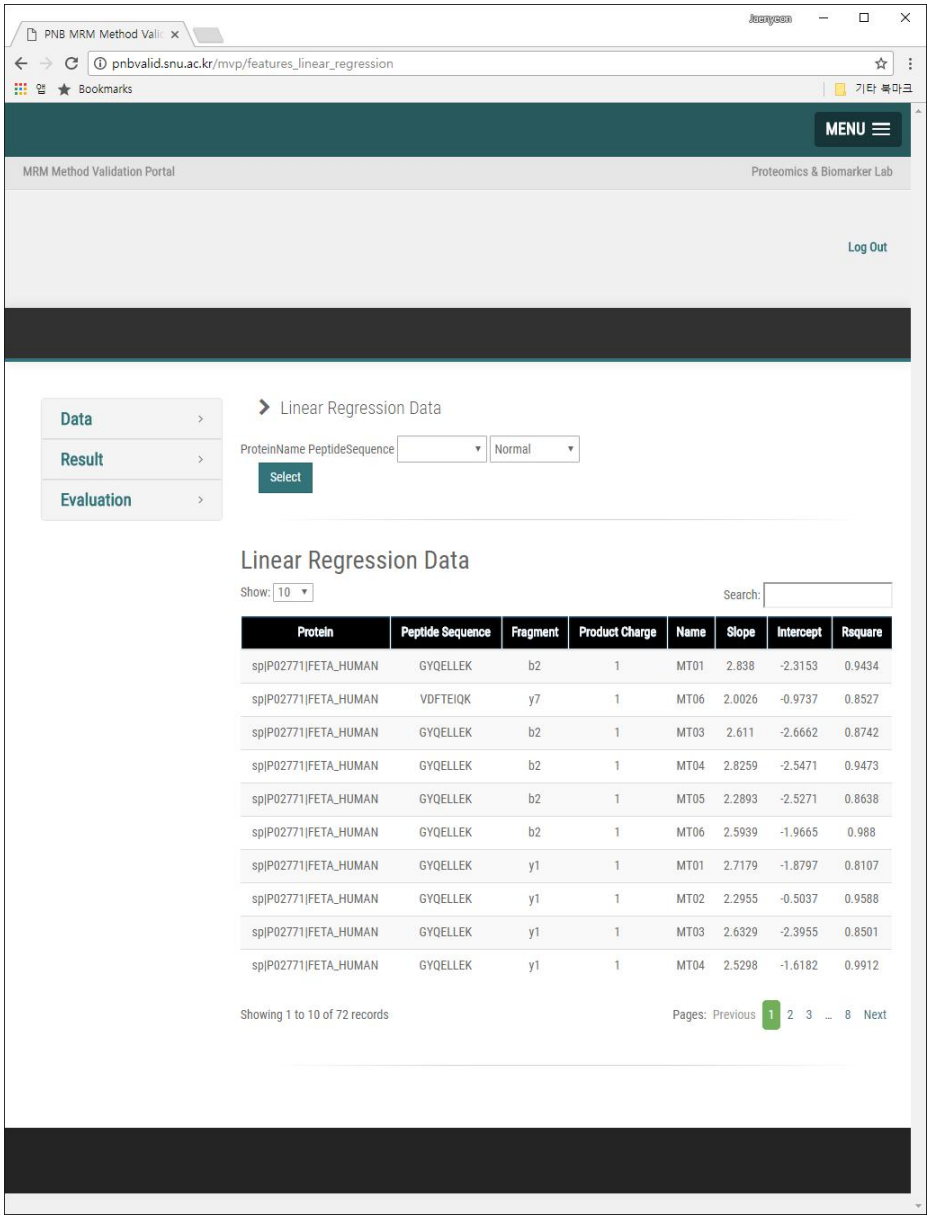


Figure 10. Calibration Curve Result Page

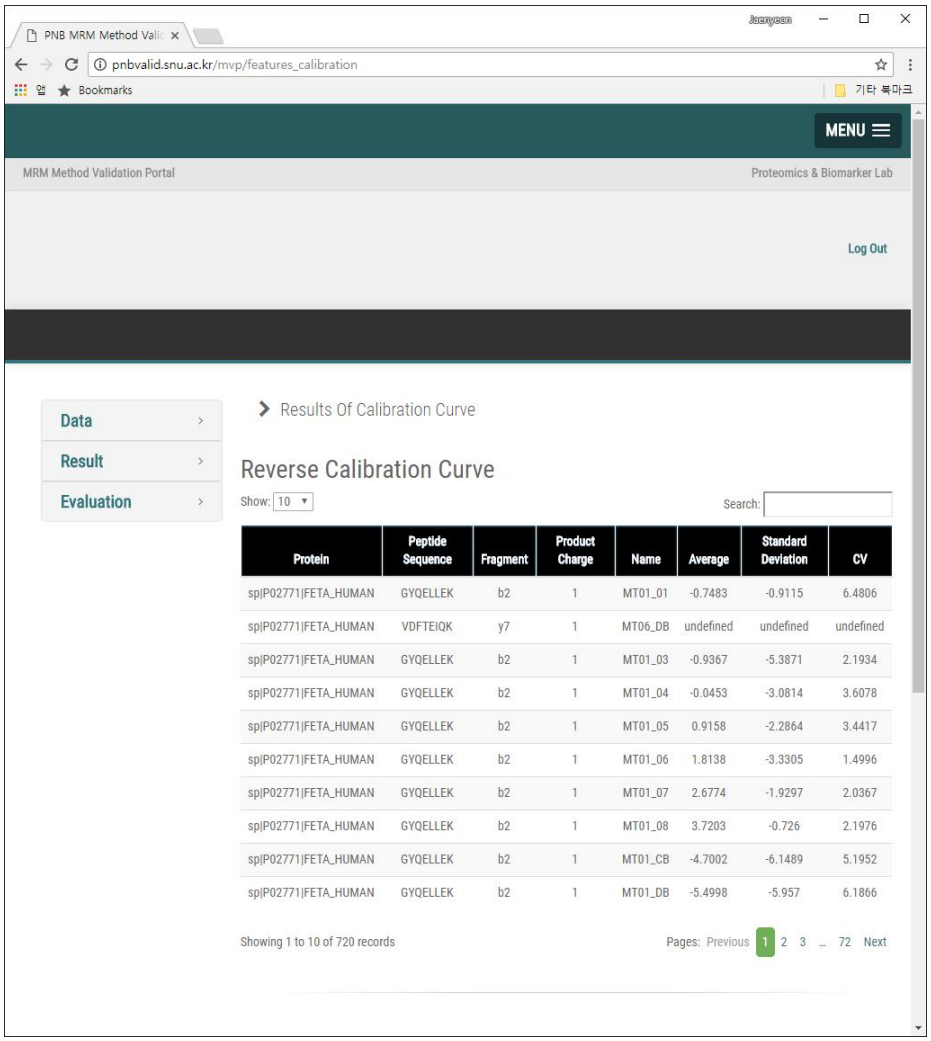


Figure 11. Selectivity Result Page

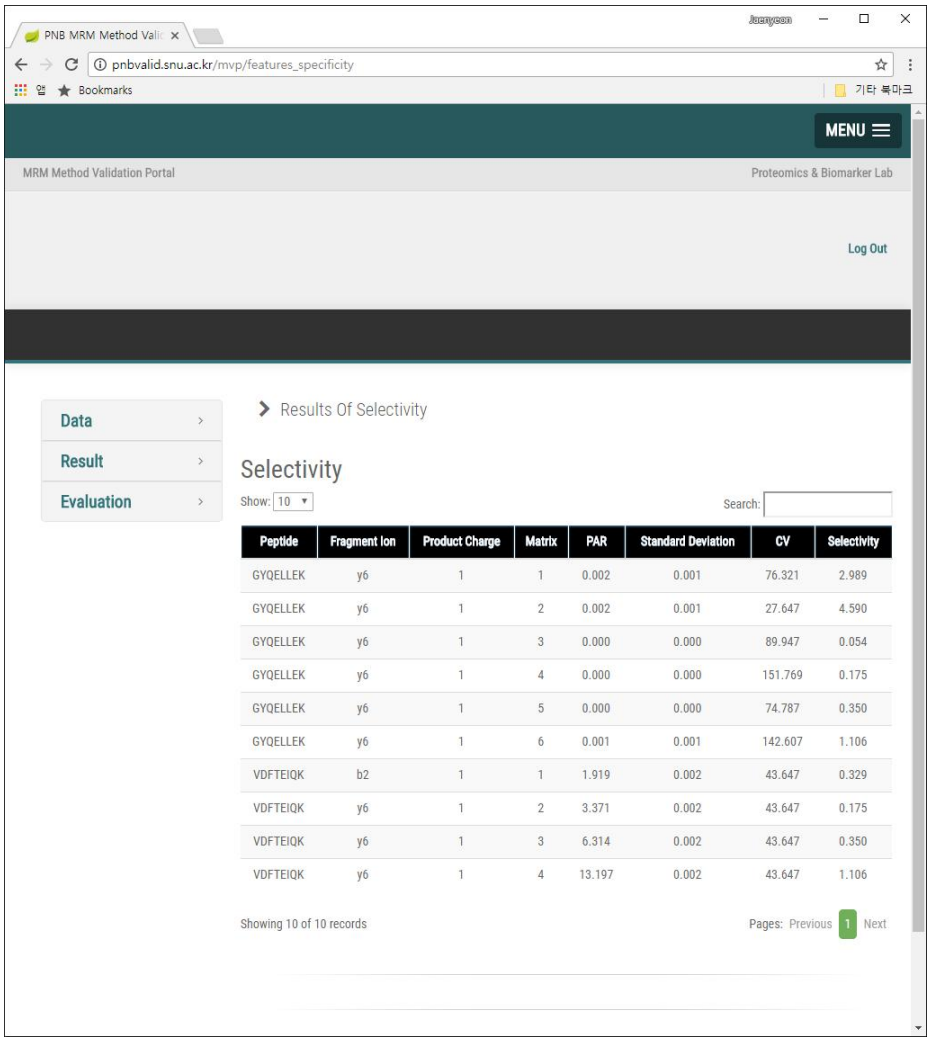


Figure 12. Sensitivity Result Page

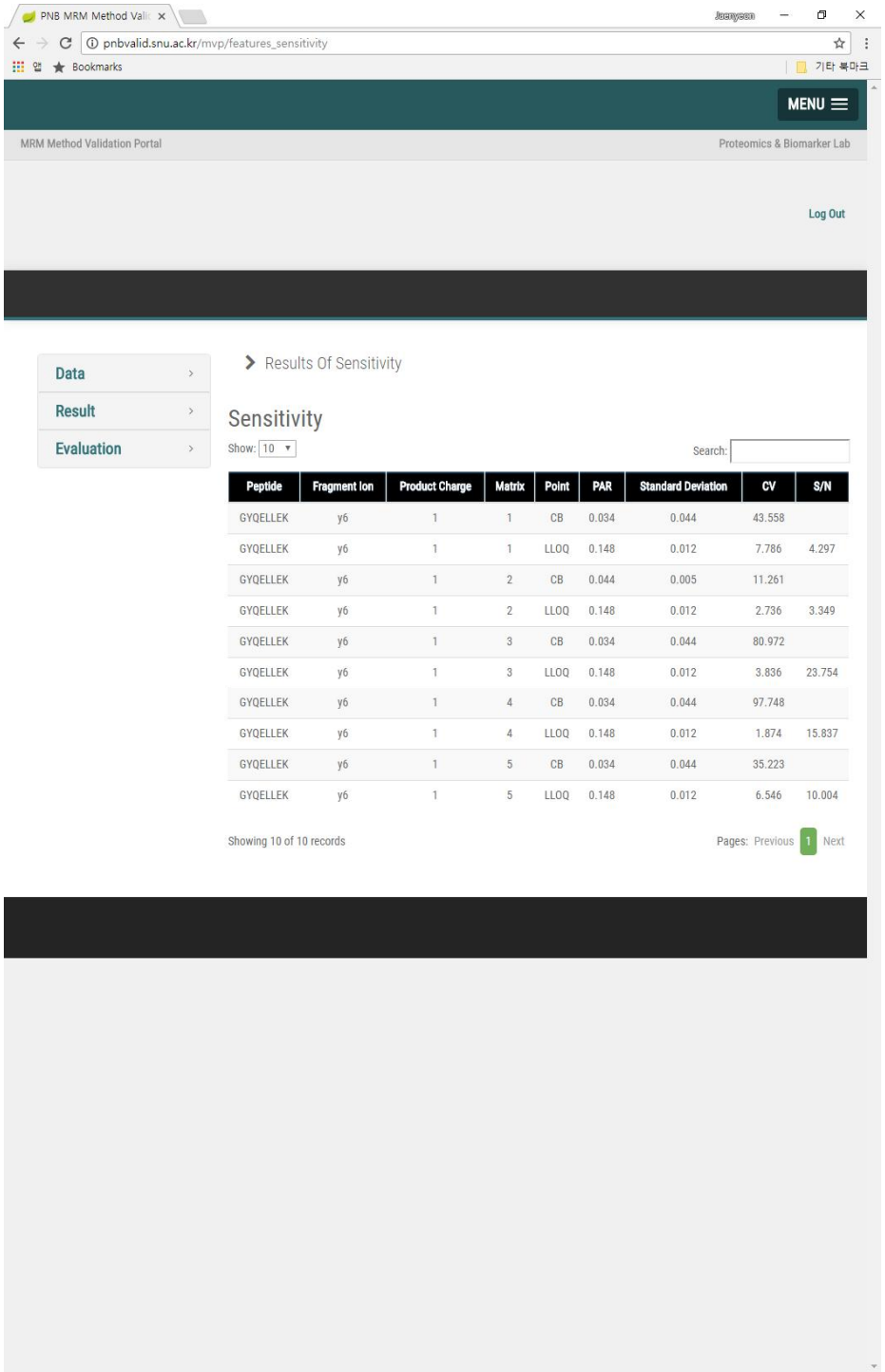


Figure 13. Carryover Result Page

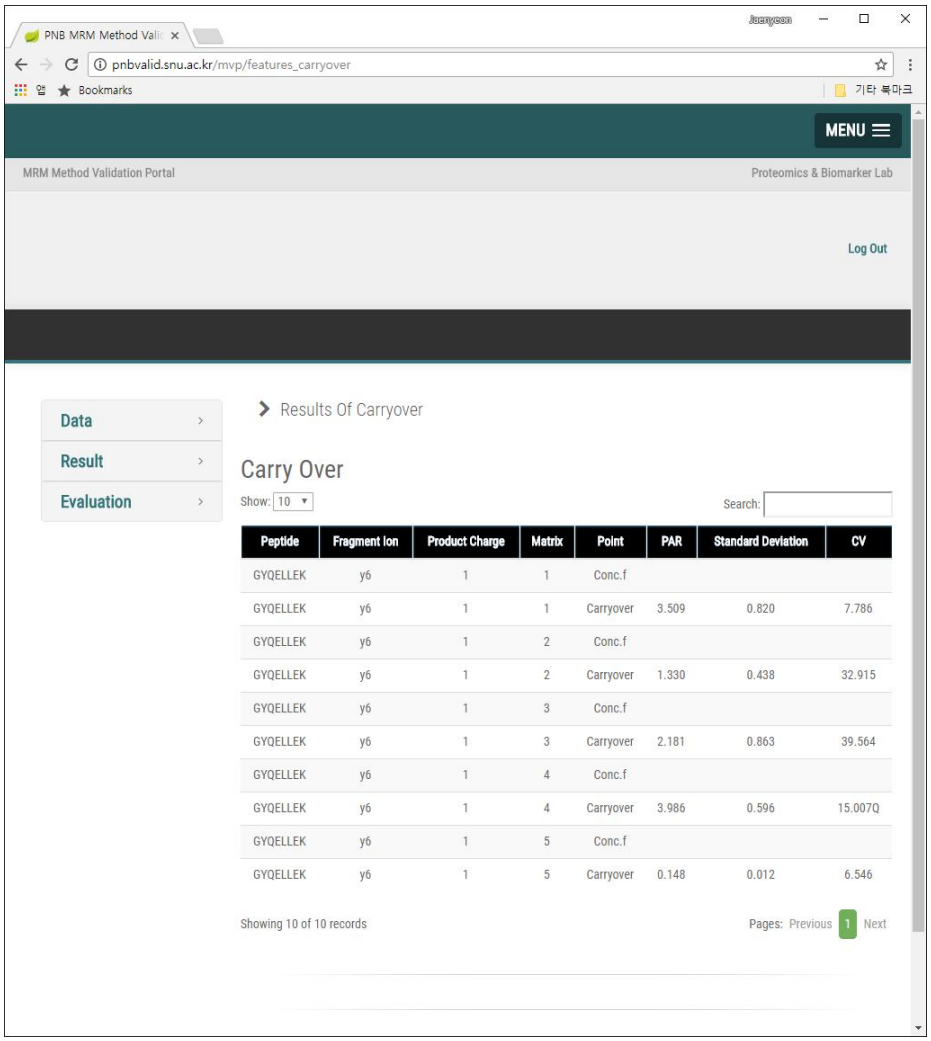


Figure 14. QC Result Page

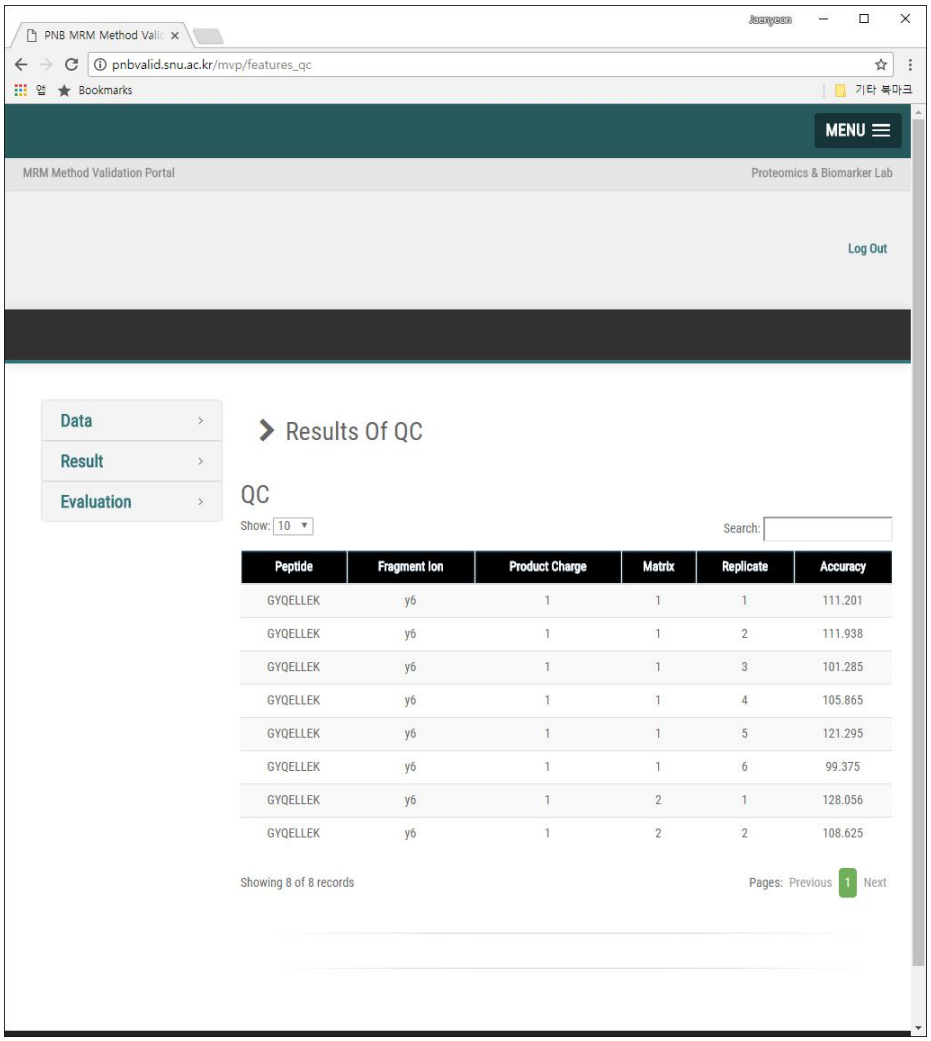


Figure 15. Precision and Accuracy Result Page

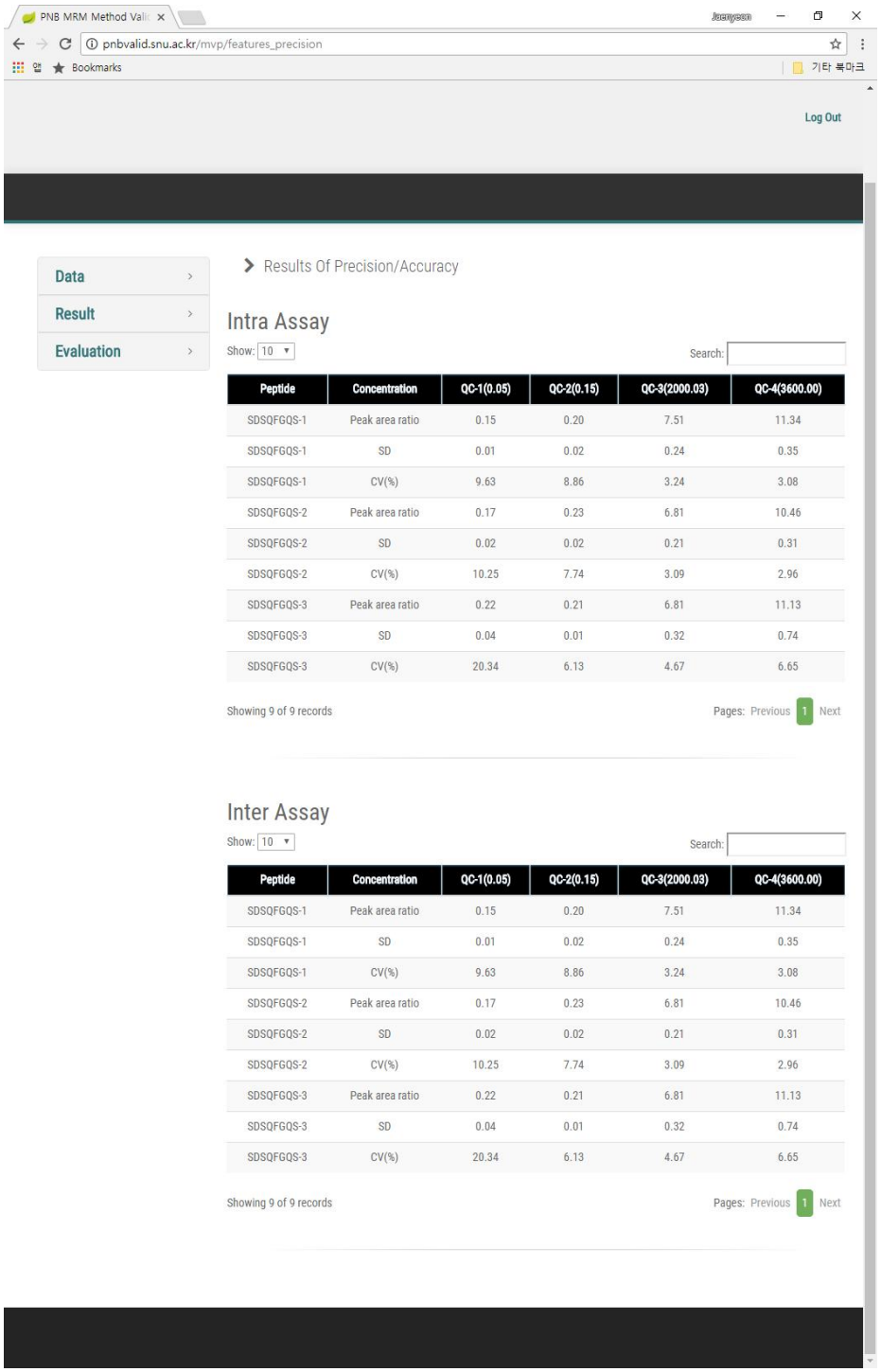


Figure 16. Matrix Effect Result Page

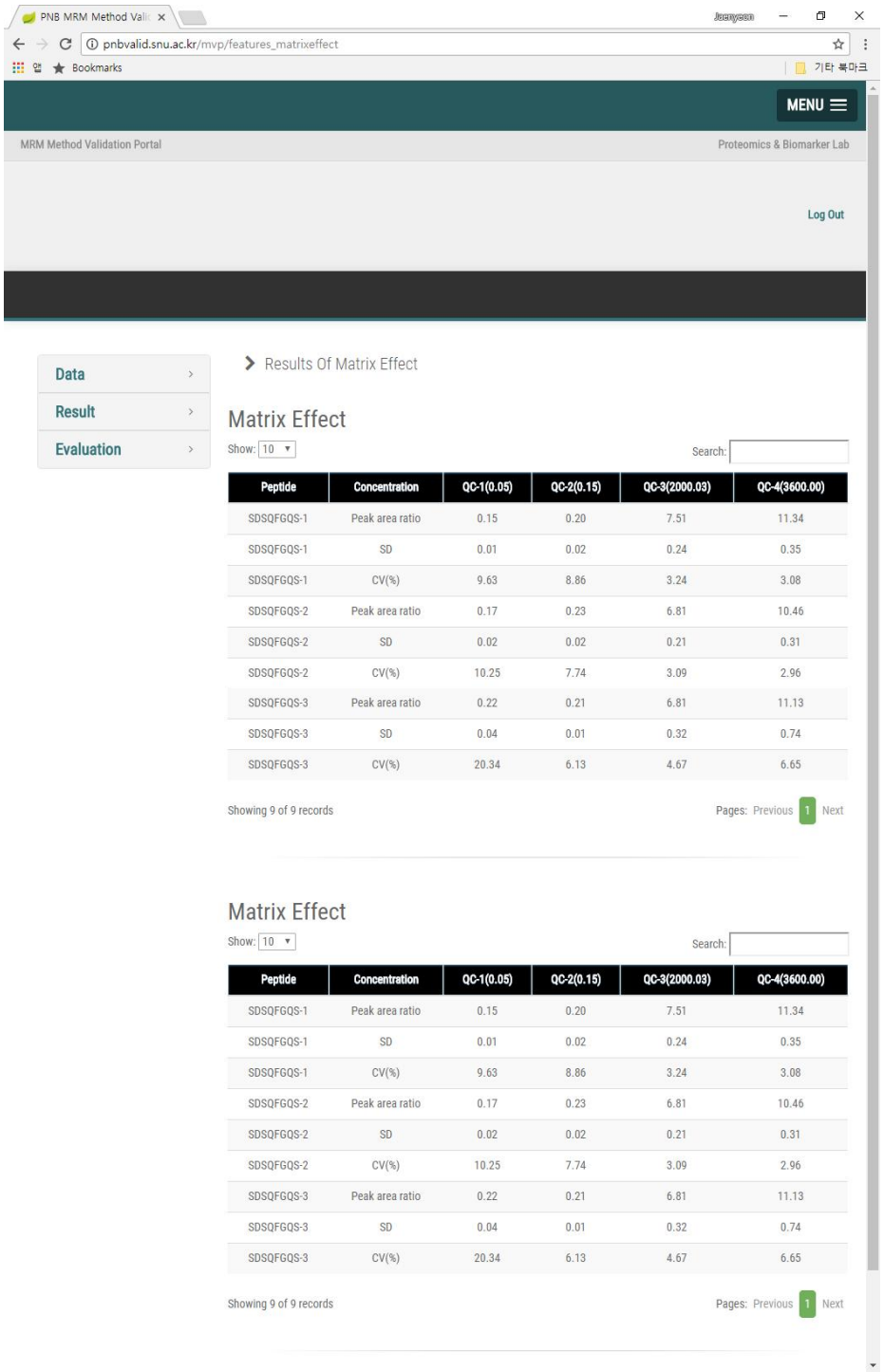


Figure 17. Recovery Result Page

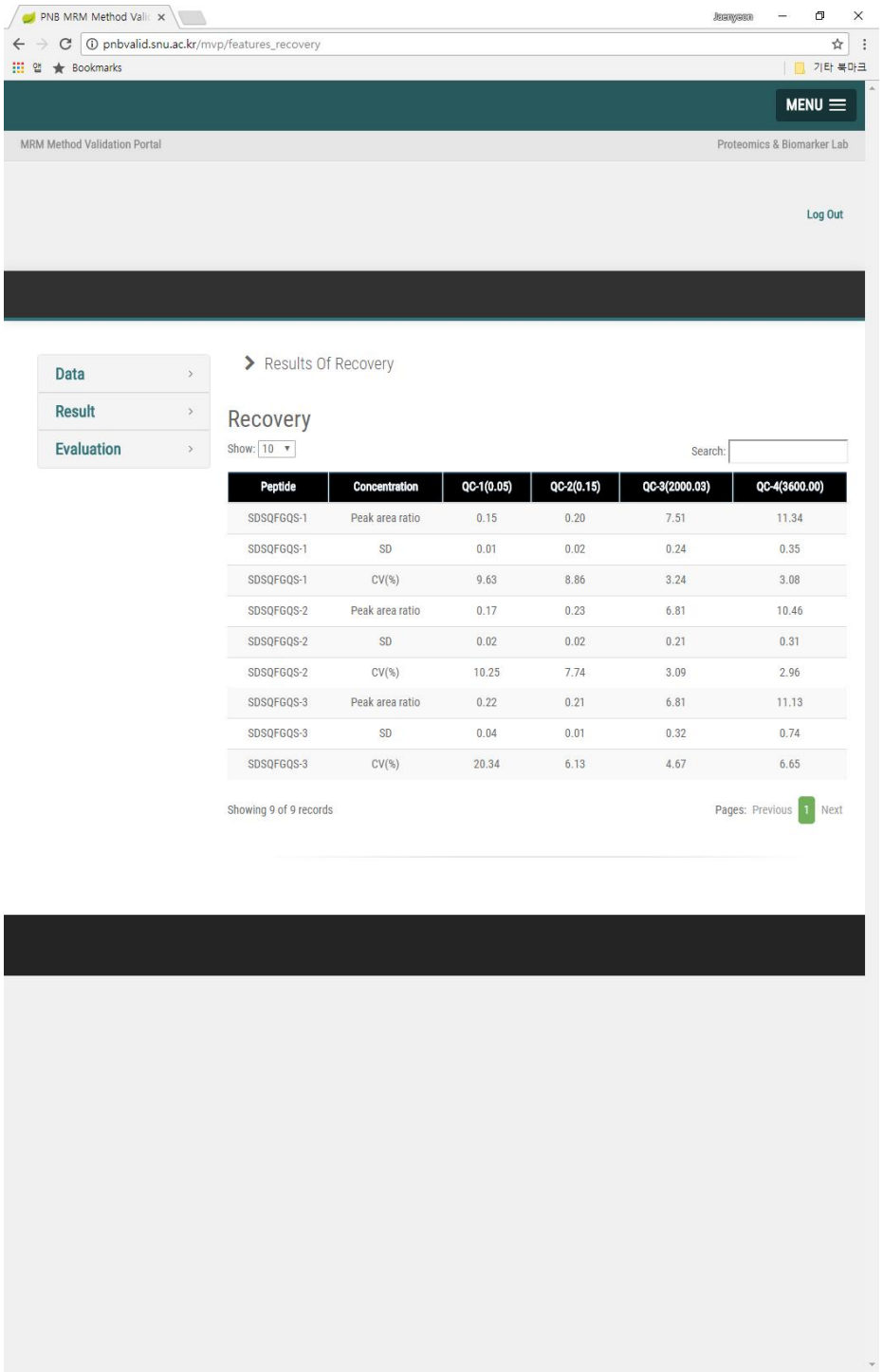


Figure 18. Stability Result Page

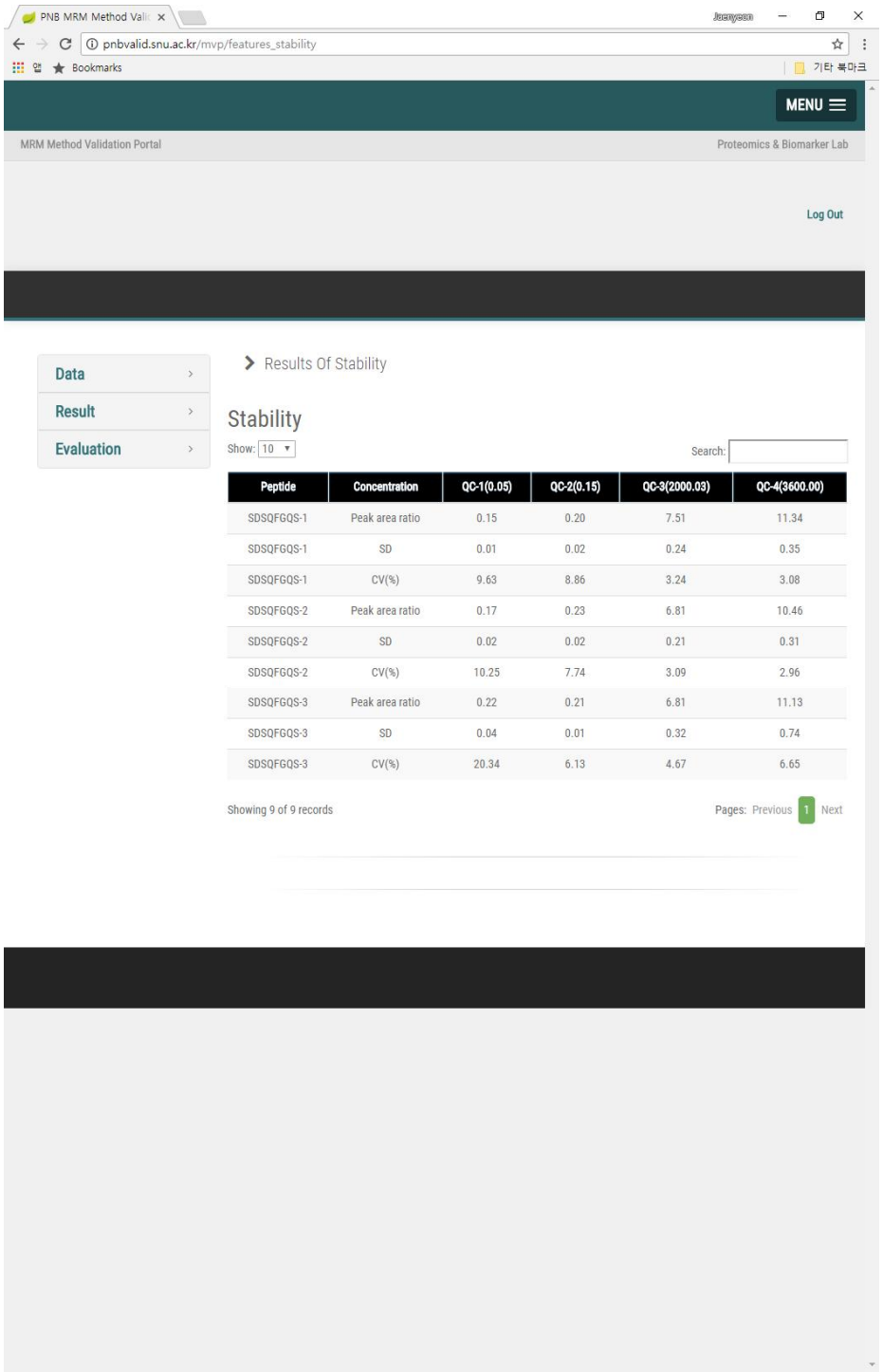
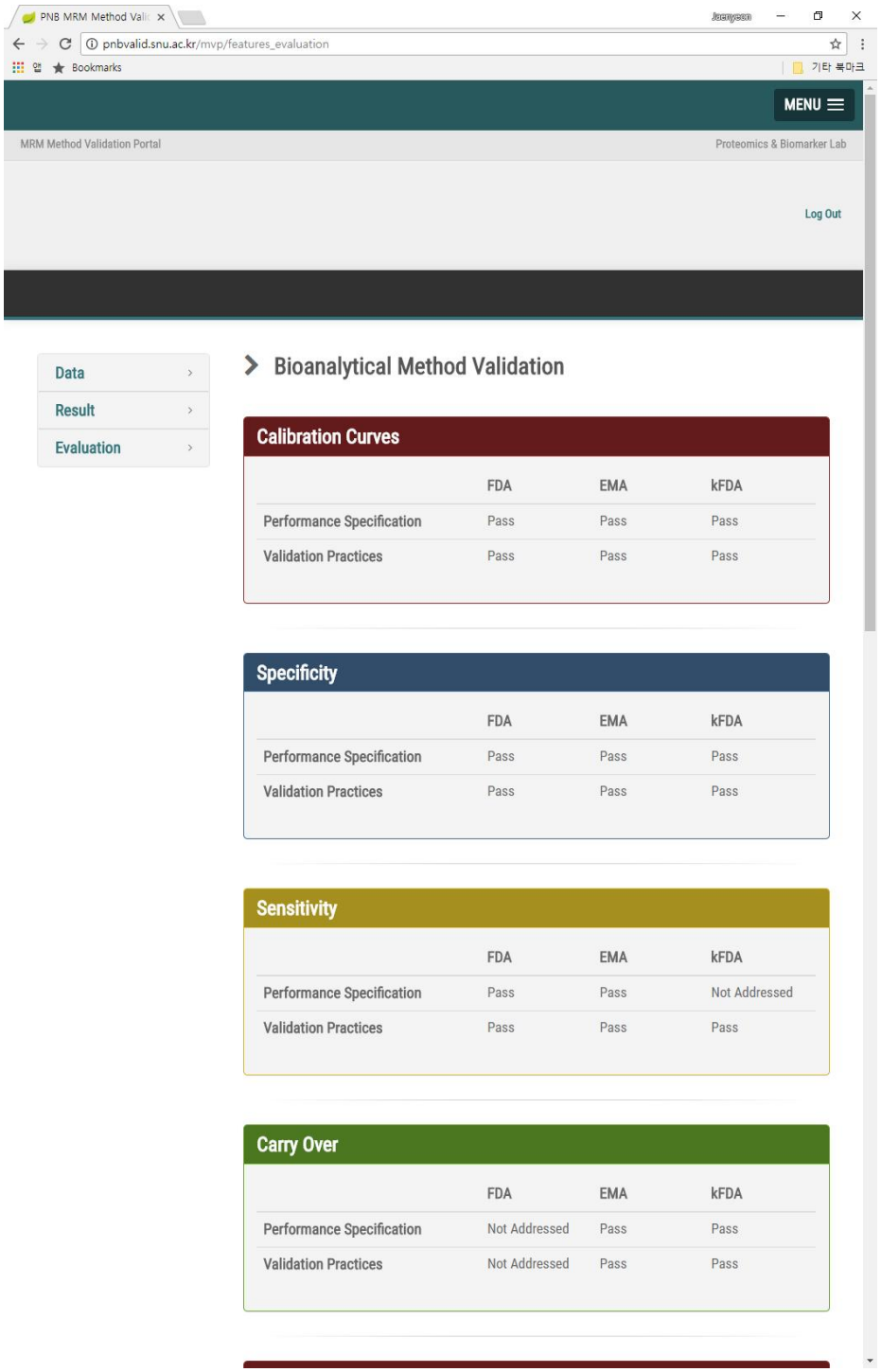


Figure 19. Evaluation Result Page



Chapter4. Discussion

We have developed and launched Portal, first web-based system to store and calculate Skyline output data for method validation. As an online Portal, calculation are done on the server side, which lowers users' computational power.

We compared data processing time between MRMMVP and several technicians, which shows that it costs lesser time and men power by using MRMMVP when there was only 2 targets. This Portal aims to support massive targets for multi marker assay validations, and its calculation effectiveness is more demonstrative if processed with large sets of data.

Several issues are still there to develop for further. MRMMVPs calculation method is sensitive to exact field and naming restrictions of Skyline generated .csv file, so technician must be cautious to his uploading file if it fits our rule. We hope to develop a more flexible naming rule by getting feedbacks from technicians of other laboratory.

Implanting it to Skyline software is also a near goal. Calculating and validating after skyline generates data file is achievable by Skylines 3rd party implants regulations. It will drain computational power from user, but it still is a fast one-click process, which ultimately lead to automation of MRM assay.

Bibliography

1. (2012) In Micheel, C. M., Nass, S. J. and Omenn, G. S. (eds.), *Evolution of Translational Omics: Lessons Learned and the Path Forward*, Washington (DC).
2. Horvath, A.R., Lord, S.J., StJohn, A., Sandberg, S., Cobbaert, C.M., Lorenz, S., Monaghan, P.J., Verhagen-Kamerbeek, W.D., Ebert, C., Bossuyt, P.M. *et al.* (2014) From biomarkers to medical tests: the changing landscape of test evaluation. *Clin Chim Acta*, **427**, 49-57.
3. Khleif, S.N., Doroshow, J.H., Hait, W.N. and Collaborative, A.-F.-N.C.B. (2010) AACR-FDA-NCI Cancer Biomarkers Collaborative consensus report: advancing the use of biomarkers in cancer drug development. *Clin Cancer Res*, **16**, 3299-3318.
4. Jennings, L., Van Deerlin, V.M., Gulley, M.L. and College of American Pathologists Molecular Pathology Resource, C. (2009) Recommended principles and practices for validating clinical molecular pathology tests. *Arch Pathol Lab Med*, **133**, 743-755.
5. Teutsch, S.M., Bradley, L.A., Palomaki, G.E., Haddow, J.E., Piper, M., Calonge, N., Dotson, W.D., Douglas, M.P., Berg, A.O. and Group, E.W. (2009) The Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Initiative: methods of the EGAPP Working Group. *Genet Med*, **11**, 3-14.
6. Kim, H., Yu, S.J., Yeo, I., Cho, Y.Y., Lee, D.H., Cho, Y., Cho, E.J., Lee, J.H., Kim, Y.J., Lee, S. *et al.* (2017) Prediction of Response to Sorafenib in Hepatocellular Carcinoma: A Putative Marker Panel by Multiple Reaction Monitoring-Mass Spectrometry (MRM-MS). *Mol Cell Proteomics*, **16**, 1312-1323.
7. Yu, S.J., Kim, H., Min, H., Sohn, A., Cho, Y.Y., Yoo, J.J., Lee, D.H., Cho, E.J., Lee, J.H., Gim, J. *et al.* (2017) Targeted Proteomics Predicts a Sustained Complete-Response after Transarterial Chemoembolization and Clinical Outcomes in Patients with Hepatocellular Carcinoma: A Prospective Cohort Study. *J Proteome Res*, **16**, 1239-1248.

초 록

Multiple reaction monitoring 질량분석법 (MRM-MS)은 질량분석 과정중 멀티플레스 어세이의 펩타이드와 관련된 트랜지션을 모니터링 하는 분석법이다. 오픈소스 툴들을 사용 및 자바를 사용하여 개발한 어세이 포탈은 FDA,EMA,KFDA에서 재시한 생체시료 분석법 밸리데이션 가이드라인에 따라 실험 결과가 검증기준의 통과여부를 확인 할 수 있는 웹사이트이다. 검증 계산된 항목으로는: 검량선, 정밀성, 정확성, 선택성, 생체시료효과, 캐리오버, 회수율, 희석 타당성, 안정성, 최저 정량 한계 항목이 있다. 어세이 포탈은 어세이 개발에 분석 워크플로우를 따라기에 에러를 일으킬 확률이 낮다. 어세이 포탈은 관리자의 계정생성에 의하여 http://pnbvalid.snu.ac.kr에서 자유롭게 접속 및 사용을 할 수 있게 개발을 하였다.

주요어 : Method Validation, Portal, Database, Biomarker, Proteomics, MRM
학 번 : 2016-21166