

Statistical Considerations of Optimal Study Design for Human Plasma Proteomics and Biomarker Discovery

Cong Zhou,^{†,‡,§} Kathryn L. Simpson,^{†,§} Lee J. Lancashire,[†] Michael J. Walker,^{†,‡} Martin J. Dawson,[†] Richard D. Unwin,[‡] Agata Rembielak,^{||} Patricia Price,^{||} Catharine West,[⊥] Caroline Dive,^{*,†} and Anthony D. Whetton^{*,‡}

[†]Clinical and Experimental Pharmacology Group, Paterson Institute for Cancer Research, Manchester Academic Health Science Centre, Christie Hospital, University of Manchester, Wilmslow Road, Manchester M20 4BX, United Kingdom

[‡]Stem Cell and Leukaemia Proteomics Laboratory, School of Cancer and Enabling Sciences, Manchester Academic Health Science Centre, Wolfson Molecular Imaging Centre, University of Manchester, M20 3LJ, United Kingdom

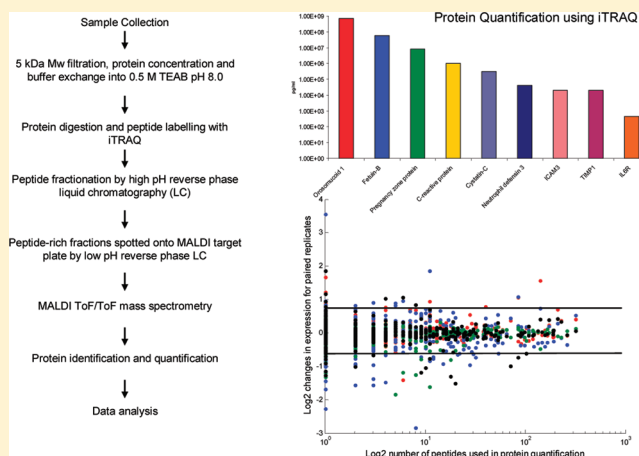
^{||}Academic Radiation Oncology, School of Cancer & Enabling Sciences, University of Manchester, Manchester Academic Health Science Centre, Christie Hospital, Wilmslow Road, Manchester, M20 4BX, United Kingdom

[⊥]Translational Radiobiology Group, School of Cancer & Enabling Sciences, University of Manchester, Manchester Academic Health Science Centre, Christie Hospital, Wilmslow Road, Manchester, M20 4BX, United Kingdom

S Supporting Information

ABSTRACT: A mass spectrometry-based plasma biomarker discovery workflow was developed to facilitate biomarker discovery. Plasma from either healthy volunteers or patients with pancreatic cancer was 8-plex iTRAQ labeled, fractionated by 2-dimensional reversed phase chromatography and subjected to MALDI ToF/ToF mass spectrometry. Data were processed using a *q*-value based statistical approach to maximize protein quantification and identification. Technical (between duplicate samples) and biological variance (between and within individuals) were calculated and power analysis was thereby enabled. An *a priori* power analysis was carried out using samples from healthy volunteers to define sample sizes required for robust biomarker identification. The result was subsequently validated with a *post hoc* power analysis using a real clinical setting involving pancreatic cancer patients. This demonstrated that six samples per group (e.g., pre- vs post-treatment) may provide sufficient statistical power for most proteins with changes >2 fold. A reference standard allowed direct comparison of protein expression changes between multiple experiments. Analysis of patient plasma prior to treatment identified 29 proteins with significant changes within individual patient. Changes in Peroxiredoxin II levels were confirmed by Western blot. This *q*-value based statistical approach in combination with reference standard samples can be applied with confidence in the design and execution of clinical studies for predictive, prognostic, and/or pharmacodynamic biomarker discovery. The power analysis provides information required prior to study initiation.

KEYWORDS: mass spectrometry, plasma proteomics, biomarkers, power analysis



INTRODUCTION

Discovery of novel biomarkers using minimally invasive approaches is increasingly required to expedite drug development in the era of mechanism-based therapeutics and patient stratification.¹ Achieving high confidence in the discovered biomarkers is a major challenge for clinical researchers, highlighted by a dearth of successful biomarker validation recently. Difficulties in validating tissue and blood borne biomarkers include the lack of availability of patients' samples, the lack of consistency in sample collection, heterogeneity in patient populations and current technological limitations. The

development of plasma biomarkers is attractive as repeat sample collection is simple and minimally invasive.¹

Human biological variation and the considerable range in specific protein concentrations within plasma present a challenge to quantitative biomarker discovery. Advances in mass spectrometry (MS)-based proteomic technologies have resulted in an increased ability to quantify and overcome such issues with careful experimental design. We have previously used an 8-channel isobaric tagging method (iTRAQ)² coupled

Received: July 6, 2011

Published: February 17, 2012

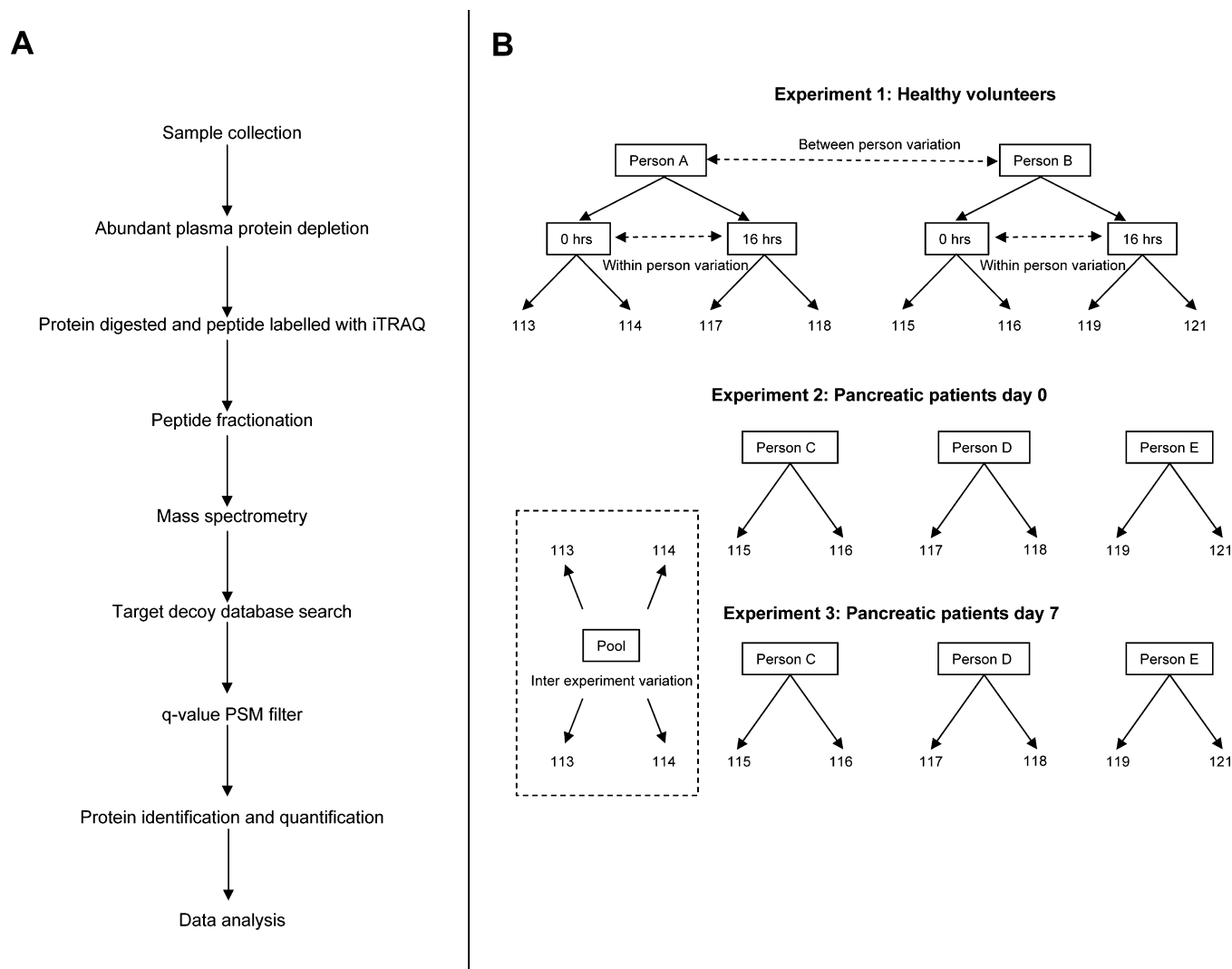


Figure 1. Design of the 8 channel isobaric tagging experiments for relative quantification of proteins from plasma. (A) Methodological workflow of sample analysis. Plasma depletion was achieved using an antibody based removal of the 20 major proteins found in human plasma. This was followed by tryptic digestion of the analyte and peptide tagging in the 8 different samples with 8 distinct isobaric tags that enable relative quantification of peptides from the 8 samples by tandem mass spectrometry. Peptides from the 8 samples were pooled and then fractionated using high pH reverse phase liquid chromatography (LC). Fractions were then spotted onto MALDI target plates by low pH RP LC and plates were analyzed by MSMS in 5800 MALDI ToF/ToF instrument (Applied Biosystems). A number of data manipulations were then performed to assess the value of the workflow. (B) Experimental setup and labeling. In Experiment 1, plasma from healthy individuals was taken 16 h apart and processed and analyzed in duplicate to assess technical and biological (within and between person) variation. In experiments 2 and 3, plasma was isolated from patients enrolled in the PACER pancreatic cancer trial at the Christie Hospital, Manchester, U.K. and 2 pretreatment samples were taken one week apart. To compare across experiments, a pooled reference control was used (pool) containing an aliquot of all plasma samples from experiments 2 and 3 points prior to sample depletion and in a 1:1 ratio for all samples.

with 2-dimensional liquid chromatography (LC) and tandem mass spectrometry (MS/MS) to quantify proteins.³ This has been shown to be a sensitive proteomic quantification method.⁴

For successful biomarker discovery, a procedure to address correctly formulated clinical research questions where power analysis is absolutely essential to experimental design is required.⁵ Furthermore, the sample sizes for such studies (MS-based or otherwise) must be feasible to allow large-scale, longitudinal clinical studies to be conducted with a high probability of identifying biomarkers with confidence. Several studies have highlighted the promise of 4 channel iTRAQ as a tool for identifying potential biomarker signatures of disease and potentially of drug response using serum, plasma, cerebrospinal fluid and tissue.^{6–10}

Isobaric tag comparisons are typically analyzed with respect to experimentally determined thresholds where a change in protein expression outside this range is deemed to be significant.^{11,12} However, false positive error (protein incorrectly determined as differentially expressed) or false negative error (protein that is truly differentially expressed not detected) can result. The power of a test is its ability to correctly lead to the rejection of the null hypothesis: the ability to detect an effect, if the effect exists. This depends on specific factors including variance in protein expression, effect size (the change in protein expression), number of replicates, and the significance level required. Therefore, to increase the power in an experiment, the number of replicates must be sufficient to distinguish between true differences and random effects. Too many replicates can be an unnecessary waste of time and

resource, whereas an underpowered study will not detect protein changes with statistical significance. The strength of including an evaluation of statistical power to enhance the experimental design of proteomic studies has been highlighted.⁵ We have extended this analysis to 8 channel relative quantification for detecting changes in plasma protein abundance. Using data sets derived from healthy controls, we have carried out a power analysis that provides us with guidance for future clinical studies. These results have been subsequently validated using data sets from pancreatic cancer patients. In addition, we have validated our method to allow interexperiment comparability via reference standards generating the first gel free proteomic approach with power analysis for direct application to clinical trials.

MATERIALS AND METHODS

Technical Workflow

An overview of the complete workflow used in this study is shown in Figure 1A.

Experimental Design

Using the workflow summarized in Figure 1A, three iTRAQ experiments were performed to meet the study objectives (Figure 1B). We first examined the technical, temporal and biological variation between samples. A positively identified protein had to meet stringent criteria based on statistical considerations using approaches designed to minimize the number of false positive identifications. Technical variability was determined by analysis of four replicate plasma samples from two healthy human controls processed independently under identical conditions and run as a single multiplexed 8 channel iTRAQ experiment (Experiment 1). These data were then used in a power calculation to evaluate the suitability of this workflow for future clinical studies.

Guided by the *a priori* power analysis from Experiment 1, the second two experiments were carried out interrogating samples derived within the PACER -TRANS substudy to the PACER clinical trial (Christie Hospital, Manchester, U.K.). PACER is a phase II study of high dose rate radiotherapy and EGFR inhibitor monoclonal antibody erbitux (Cetuximab) in patients with locally advanced pancreatic cancer. Prior to treatment, blood samples were collected from patients at two different time points with one week apart (day 0 and day 7, Experiments 2 and 3; Figure 1B). The objective is to identify proteins that are differentially expressed in the 7 day period and carry out a *post hoc* power analysis which allowed us to explore the validity of the *a priori* power calculation using Experiment 1. Furthermore the study design will address use of a pooled reference sample, created by mixing part of each sample used in Experiments 2 and 3. This would allow the direct comparison of protein changes determined from different 8-plex experimental runs, essential for the future of large scale trial analyses conducted by this method.

Human Plasma Samples

Blood was collected from donors in lithium heparin coated tubes (BD Vacutainer) and centrifuged within 30 min of collection at 2500× *g* for 15 min at 4 °C before aliquots of the plasma layer were stored at −80 °C. Samples were collected at two different time points for each patient and healthy volunteer. For healthy volunteers samples were collected 16 h apart. Blood samples were taken from 3 patients with pancreatic cancer enrolled in the PACER study at the Christie

Hospital, Manchester, UK (ref.06/Q1407/17) following written informed consent with ethical approval from the Central Manchester Local Research Ethics Committee. Two blood samples were taken one week apart, prior to patients receiving any therapy. Pooled samples were created prior to depletion by the accumulation of 50 μL of each plasma sample from all three pancreatic patients at both time points (Figure 1B).

Protein Depletion, Digestion and Labeling

Abundant proteins were removed from plasma using a Sigma Top20 spin column following the manufacturers' protocol (Sigma Aldrich). Depleted samples were concentrated and buffer-exchanged into 1 M TEAB using Vivaspin 500 centrifugal concentrators (Sigma Aldrich) as per manufacturer's instructions. The protein concentration in buffer-exchanged samples was measured using the 2-D Quant kit (Amersham Bioscience, Buckinghamshire). Fifty μg of each sample was reduced with the addition of 1/10th of the sample volume of 50 mM *tris*(2-carboxyethyl)phosphine for 1 h at 60 °C. Cysteine residues were then alkylated by the addition of 1/20th of the total sample volume 200 mM Methyl thiomethanesulfonate (in isopropanol) before incubation for 10 min at room temperature. Protein was digested by the addition of 5 μg of porcine trypsin (Sigma Aldrich) with 15 min in a CEM discoverer microwave at 55 °C (CEM, North Carolina) to aid digestion, followed by overnight at 37 °C. The digested protein samples were labeled with 8plex iTRAQ reagents according to the manufacturers' instructions (Applied Biosystems, Foster City, CA). After labeling the samples were dried at 60 °C in a SpeedVac and then stored at −20 °C. Samples were labeled according to Figure 1B.

High pH Reverse Phase (RP) Chromatography

iTRAQ labeled samples were reconstituted in 100 μL of 0.1% Ammonium hydroxide (Solvent A) and pooled prior to being loaded onto a 100 × 4.6 mm 3 μm C18 HPLC column (Fortis, Cheshire, UK). Peptides were eluted by the application of a linear 30 min gradient up to 50% solvent B (Acetonitrile, 0.1% Ammonium hydroxide) with 70 × 15 s fractions collected from 4 min. Fractions were dried in a SpeedVac at 60 °C and stored at −20 °C.

Liquid Chromatography (LC)

Dried samples were reconstituted in 130 μL of 0.1% TFA, 2% ACN. Half of the sample was loaded onto a trap column using a U3000 liquid chromatography system (Dionex, Sunnyvale, CA) and the peptides fractionated by a capillary RP C18 HPLC column (Acclaim PepMap C18, 3 μm 100 Å) at a flow rate of 0.8 μL/min with a gradient of between 2 and 40% acetonitrile, 0.1% TFA. The flow-through was spotted onto a MALDI plate (AB SCIEX, Foster City, CA) in 15 s fractions using an online Probot (Dionex, Sunnyvale, CA) with α -cyano-4-hydroxycinnamic acid mixing with the eluent to a final concentration of 1.25 mg/mL.

Mass Spectrometry (MS/MS)

Mass spectrometry was carried out on an AB Sciex TOF/TOF 5800 (AB Sciex, Foster City, CA, USA) using 1000 shots for MS. MS/MS was carried out on the top 27 precursors with a S/N of higher than 8 using 4000 laser shots, a 2Kv acceleration voltage and air as the collision gas. MS/MS spectra were smoothed using the Savitsky Golay Algorithm with 3 points and 4 orders of magnitude.

Protein Database Searching

All MS/MS data were submitted to ProteinPilot software version 3.0 (Applied Biosystems) for database searching and iTRAQ reporter ion quantification. Searches were performed against the IPI Human (v3.59) protein sequence database, containing 160248 protein sequences. A reversed database was searched at the same time to control the false discovery rate (FDR) of protein identification (see below). Cys alkylation with methanethiosulfate (MMTS) and trypsin as the digestion enzyme were specified in the search. Biological modifications and amino acid substitutions were also permitted. ProteinPilot uses the Pro Group Algorithm to ensure that any peptide ID is only represented by one protein ID.

False Discovery Rate of Protein Identification

The FDR of protein identification was calculated using a target-decoy searching strategy¹³ where forward and reverse sequences from the database were in equal competition to be the highest ranking identification for each spectrum. The q -value¹⁴ approach was then used to define a peptide confidence threshold at which to call PSMs significant as to minimize false positives. The protein level FDR was estimated using the method reported by Kall et al.¹⁵ The maximum allowed peptide FDR and protein FDR are set to 1% and 5% respectively.

Protein Quantification

Peptides with no quantification, absence of one or more reporter ions, low signal-to-noise ratio or with confidence <1% were not used. If peptides were only partially enzymically hydrolyzed, missing an iTRAQ reagent label, or contained a low probability modification then they were also removed. Additionally, peptides shared among related, but distinct proteins or peptides where the spectrum is also matched to a different protein with unrelated peptide sequence were not used in quantification. Remaining peptides were included as contributing factors to protein quantification. Further, if a protein contained no peptides above the peptide confidence threshold determined by q -value analysis, it was judged to have failed identification and quantification and subsequently excluded from the final data set. Protein quantification was then calculated manually as per ProteinPilot software:

$$\text{Weighted average of log ratio} = \frac{\sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i}$$

where x_i is the log(peptide ratio) for the i th observation, w_i is the weight for the i th observation normalized against the percentage error under the peak to remove biases cause by label differences. Finally, n is the number of contributing peptides to a protein's average ratio.

MS Variation

The unweighted standard deviation (Std) of each protein ratio was calculated using the following equation from ProteinPilot:

$$S = \left[\frac{\sum_{i=1}^N (x_i - x_{\text{avg}})^2}{n - 1} \right]^{0.5}$$

where x_i is the log(peptide ratio) for the i th observation, x_{avg} is the unweighted average of x_i and n is the number of peptide ratios contributing to a protein's average ratio. The average of all Stds calculated from each protein identified and quantified was then used as an estimate of the MS variation.

Sample Size Determination

Sample size calculations were based on the normal linear mixed effects model as described previously.^{16–18} The \log_2 ratio represented the ratio change between the iTRAQ labels. The effect size was calculated as follows, where rep1 refers to replicate 1 and rep2 refers to replicate 2.¹⁹

$$\text{effect size} = \log_2 \left(\frac{\text{rep1}}{\text{rep2}} \right)$$

For example, for a 2 fold change, the effect size = $\log_2(2) = 1$. Therefore the null hypothesis is:

$$\text{effect size} = 0$$

We will accept or reject this hypothesis according to the observed experimental data.

The approach utilized by Dobbin and Simon²⁰ for sample size calculations in microarrays was adopted such that the \log_2 ratio of each protein p had variance across all samples within a group of interest composed of both technical (σ_p^2) and biological (τ_p^2) variance. In a two-group problem (e.g., pre- vs post-treatment) the total number of biologically distinct samples n in each class is given by:

$$n = 2 \left[\frac{z_{\alpha/2} + z_{\beta}}{\delta} \right]^2 \left(\tau_p^2 + \frac{\sigma_p^2}{m} \right)$$

where m is the number of technical replicates per sample, δ is the difference in class means or observed effect size. $z_{\alpha/2}$ and z_{β} are the 100 $\alpha/2$ th and 100 β th percentiles of the normal distribution. These are specified by the significance level α and the power $1 - \beta$ that we wish to base our hypothesis around.

Technical variance (σ_p^2) was estimated from four replicate plasma samples processed identically and run as a single iTRAQ MS/MS experiment. This allowed for an assessment of variation caused by the experimental workflow, where ideally ratios of all the proteins quantified should be 1. The technical replicates were 114:113, 116:115, 118:117 and 121:119.

Biological variation (τ_p^2) comprises within person variation and between person variation. Using technical variation and within person variation, proteins that are differentially expressed within a specific patient can be derived. With the additional between person variation, proteins that are differentially expressed in all patients can be derived. These proteins can be used as candidates for biomarkers. In Experiment 1, within person variance was estimated using the within person variation across a 16 h time period, and between person variance was estimated using the between person variation among the two healthy donor controls. Any deviations from a ratio of 1 would provide information regarding natural variation. Of course, the observed biological variation naturally contains the component from technical variation which should be excluded before power analysis.

iTRAQ Workflow Reproducibility

The same pooled reference was used for both PACER 0 day and PACER 7 day experiments to allow for direct comparison of protein ratios across two separate runs of the complete workflow. iTRAQ ratios for the reference labels in proteins quantified in both experiments were compared using the Bland-Altman comparison, and assessed statistically with Pitman's test of difference in variance (Stata 10.1, StatCorp LP).

Table 1. Use of a Target-Decoy Database Search of the Experiment 1 Data Set Using Different *q*-Value Thresholds^a

<i>q</i> -value threshold	PSMs			Quantified Peptides			Quantified Proteins		
	Target	Decoy	FDR	Target	Decoy	FDR	Target	Decoy	FDR
None	69040	16266	0.24	7947	56	0.007	462	31	0.067
0.05	30672	1614	0.05	7763	53	0.007	459	29	0.063
0.01	21900	230	0.01	6918	26	0.004	428	12	0.028
0.001	18828	77	0.002	6208	2	0.0003	391	2	0.005

^aThe number of target and decoy Peptide Spectral Matches (PSMs), quantified peptides and quantified proteins are shown for four choices of *q*-value threshold together with the calculated FDR (no threshold indicates the ProteinPilot default output).

Western Blotting

One μg of undepleted plasma was diluted 10-fold in 10 mM phosphate buffered saline (PBS), followed by the addition of 2 \times Laemmli buffer (Bio-Rad Laboratories, Hemel Hempstead, U.K.) and heated at 95 °C for 20 min prior to SDS-PAGE in 10% polyacrylamide gels. Proteins were transferred onto PVDF membranes (Perkin-Elmer, Waltham, MA), incubated in 1% (w/v) Non-Fat Milk in 10 mM PBS-Tween(T) (0.1% w/v) followed by incubation with either mouse Anti-Peroxiredoxin II 1:3000 in 1% (w/v) Non-Fat Milk in PBS-T (1E8 Ab Frontier, Korea) or Rabbit Anti-Coagulation Factor XIII B Chain Precursor (F13B) 1:1000 (HPA003827 Sigma Prestige Antibodies, St Louis, MO) and a horseradish peroxidase-coupled antimouse or antirabbit secondary IgG (Dako, Glostrup, Denmark). This was followed by detection with the Western Lightning Chemiluminescence Reagent Plus (Perkin-Elmer).

RESULTS

Parameters for Protein Identification in 8 Channel Plasma Proteomics

Experiment 1 involved samples from healthy donor controls (Figure 1B) and in this experiment 85306 mass spectra were matched after simultaneous searching of proteins against the International Protein Index reversed target decoy database, resulting in a peptide FDR of 24%. Within these peptides, 8003 spectra were quantifiable which resulted in 493 nonredundant proteins. The inclusion of low confidence peptides in protein identification/quantification led to an FDR of 6.7%. In such experiments there is a need to control for the FDR, thus a *q*-value approach was implemented²¹ (see Methods) whereby peptides were filtered based upon different confidence thresholds prior to their use in identification and quantification, thereby low confidence peptides could be excluded from further analyses and the FDR could be set at an appropriate value (Table 1).

It was evident that the use of different *q*-value thresholds varied the number of matches selected as significant (Supplementary Figure 1, Supporting Information). We found that a minimum peptide confidence of 91% was required to ensure that the false positive proportion of significant peptide spectral matches (PSMs) was <0.01 after correction for multiple testing (Table 1). A single 8 channel isobaric-tagged peptide identified with above 91% confidence was thus shown to be evidence for protein identification and quantification. Those proteins identified with no quantified peptides with $\geq 91\%$ confidence were excluded from the final data set. The peptides and proteins identified in this study are listed in Supplementary Table 1 (Supporting Information). In Experiment 1 (healthy volunteers) using the above criteria, 428 proteins were identified and successfully quantified with a protein FDR of 2.8% using the method of Käll et al.,¹⁵ 284 of

these proteins were identified with no less than 2 peptides (Supplementary Table 1). Proteins quantified with one peptide generally have a larger variance than those with more than one peptide but we observed no statistically significant difference (Welch's test $p = 0.08$).

The same *q*-value strategy was applied to Experiments 2 and 3 (pancreatic cancer patient samples) in order to minimize the FDR for protein identification after searching against a target-decoy database. In Experiment 2 (day 0), 396 (2.6% FDR) and Experiment 3 (day 7), 374 (2.8% FDR) proteins were identified and quantified (Supplementary Table 1, Supporting Information). Some of these proteins have been observed in the literature to span over 6 orders of magnitude in plasma protein concentration (Figure 1A and Methods), including intracellular, low abundance proteins such as fructose-bisphosphate aldolase B, a cytosolic protein and Interleukin 6 receptor (IL6-R). As an example of our proteomic penetration, IL6-R has been recorded at a concentration of 453 pg/mL in serum (about 9 pM).²² Supplementary Table 2 (and references therein, Supporting Information) highlights examples of some of the proteins identified together with their relative abundance in plasma.

Biological and Technical Variance in 8 Channel Isobaric Tagging Plasma Proteomics

We next sought to understand the bias caused by technical and biological variation.²³ The need for a robust statistical design at each stage of analysis in quantitative proteomic profiling experiments is paramount. Technical variation was addressed by analysis of duplicate samples prepared from healthy controls (Experiment 1, Figure 1B). We showed a high correspondence and statistically significant correlation between all technical replicate labels in this study ($p < 0.0001$) as summarized in Supplementary Figure 2 and Supplementary Table 3 (Supporting Information).

The distribution of technical variation is illustrated in Supplementary Figure 3A (Supporting Information). It can be said to resemble a Gaussian distribution but has heavier tailing. As proposed by Breitwieser et al.,²⁴ technical variation of iTRAQ data can be modeled as a Cauchy distribution. In this study, however, it was found that Cauchy distribution did not provide satisfactory fitting therefore a Gaussian approximation was carried out on truncated data (see below for detail). The accuracy and amount of data that fell within an acceptable error range contained in the technical replicates are summarized in Table 2. Here all the protein ratios were \log_2 transformed and the data were categorized into groups similar to those proposed by Gan et al.,²⁵ with variation cut-offs between 0 and 100% of the expression data. In order to estimate the technical variance of the data, a Gaussian approximation was made to the distribution of technical variation. The approximation removed the largest and smallest 1% protein ratios and fitted the

Table 2. Number of Proteins Identified (% of Total) with Different Variation Cut-offs for Technical and Biological Replicates^a

variation of log ₂ protein ratio	technical replicate (% of total)	within person (% of total)	between person (% of total)
±10%	37	63	53
±20%	61	69	61
±30%	74	75	69
±40%	83	81	74
±50%	88	84	79
±60%	92	87	82
±70%	94	88	85
±80%	96	90	87
±90%	97	92	89
±100%	98	94	90

^aUsing data obtained in Experiment 1 the accuracy and amount of data that fell within various error ranges was calculated. Between person and within person variation listed here were derived from the observed data by removing the variance component from technical variation.

remaining data using a Gaussian distribution model. A Std of 0.3 was observed representing a 95% confidence interval of ±0.59 for technical variation in log space. This variance level will be used for sample size calculations.

In contrast to technical variation, biological variation is protein, patient and disease dependent. The distributions of within- and between-person variations in Experiment 1 are illustrated in Supplementary Figure 3B and C (Supporting Information). These distributions were clearly asymmetric and can be challenging to model with existing theoretical distributions. In this study, the biological variance was calculated as a spread of typical variation values such as the 70th percentile, 85th percentile and maximum variation seen in the biological replicates, similar to what has been proposed by Yang and Speed.²⁶ In doing this, the observed within- and between-person variation were categorized using the same method as described for the technical replicates (Table 2). Greater variation was observed between person A and person B in the study than within each individual at two different time-points. It was clear that the expression level of the majority of the proteins (~80%) varied only to a limited extent.

Sample sizes for clinical proteomic trial design were then calculated using $\alpha = 0.05$ and $1-\beta = 0.8/0.7$, which represent the common choices for significance and power analyses. The effect sizes (changes in protein abundance) were taken as $\log_2(1.7)$ and $\log_2(2)$. These values are chosen to represent possible fold change cut-offs in proteomic studies from previous cell line-based studies. The technical variance and the biological variance were calculated using the method described above. The number of technical replicates of interest in each class was 1, 2, and 4. Required sample sizes were calculated according to the equation described in the Methods and the results are summarized in Table 3. It is clear that the variation of protein quantities has a dramatic effect on the number of patients that would be required in each group to adequately power a study. For example, for an experiment with 2 technical replicates per patient and a minimum required power of 0.8, 5 patients were required to consider a 1.7 fold change to be significant for proteins with variances not exceeding 70th percentile (70% least variant proteins). The required patient number increases to 14 for 80% proteins (least variant) and rises dramatically to 575 to cover all proteins. Observing a larger change in protein abundance, having more technical replicates per patient or reducing the power required for the study would allow a smaller number of patients to be required. In many clinical trials, as little as 3 patients per cohort have been recruited. According to our calculation, this would be sufficient to detect a 2 fold change with a power of 0.8 for 70% of proteins (70% least variant). However, if a study is required to detect more variant proteins, clearly at least 6 patients per cohort would be beneficial.

Application of Acquired Power Analysis Data to Material Gathered in a Clinical Trial

To test the validity of our method we applied our workflow to samples from a clinical trial in which two 'baseline' pretreatment samples from 3 patients with pancreatic cancer were taken one week apart and analyzed over two iTRAQ experiments (Figure 1B, Experiment 2, day 0 samples; Experiment 3, day 7 samples). According to our *a priori* power calculation using healthy volunteers, the patient group size would allow us to detect 2-fold changes with a power of 0.8 for 70% of proteins (least variant). Thus in Experiment 2 and 3, we aimed to verify this with a *post hoc* power analysis using

Table 3. Estimated Sample Sizes Required Per Group^a

effect size	1- β	number of replicates	variance				
			70th percentile	75th percentile	80th percentile	85th percentile	maximum
$\log_2(1.7)$	0.7	1	5	8	12	17	453
		2	4	7	11	16	452
		4	4	6	10	15	452
$\log_2(1.7)$	0.8	1	7	10	15	21	576
		2	5	9	14	20	575
		4	5	8	13	19	574
$\log_2(2.0)$	0.7	1	3	5	7	10	266
		2	3	4	6	10	265
		4	2	4	6	9	264
$\log_2(2.0)$	0.8	1	4	6	9	13	338
		2	3	5	8	12	337
		4	3	5	8	12	337

^aUsing variance data obtained from Experiment 1 (healthy volunteers) sample sizes are reported for several choices of effect size and variance level. Sample sizes required are per group. To see an effect size greater than a 2 fold change, 3 samples per group (for e.g. 3 pre-treatment vs. 3 post-treatment) with 2 technical replicates would be sufficient for proteins with 70% variance with a power of 0.8.

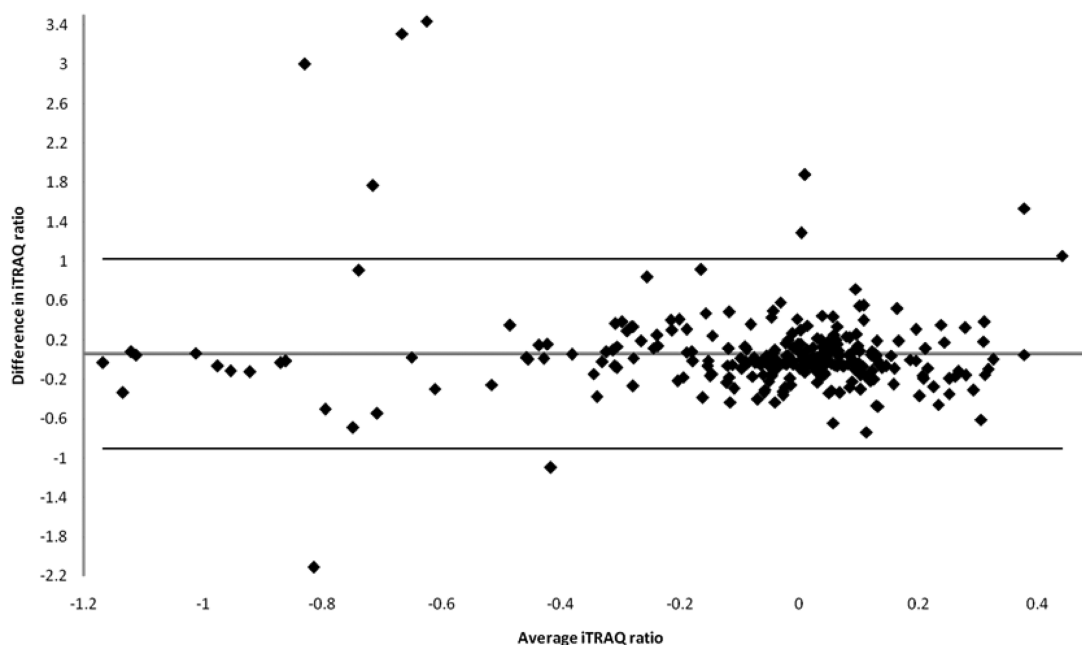


Figure 2. Bland-Altman plot for pooled reference reproducibility across iTRAQ experiments 2 and 3 (PACER day 0 and PACER day 7 clinical samples). Total of 3 patients, each with duplicate sample at day 0 and day 7 contributing equally to a pooled reference of 12 samples. Proteins that were identified and quantified in experiments 2 and 3 and originating from the pooled reference sample as defined by the incorporation of iTRAQ labels 113 and 114 were analyzed for the agreement in \log_2 iTRAQ ratios (ideally equivalent between both experiments). The agreement was calculated by the mean of the two measurements versus the difference in values, thus the smaller the difference the greater the reproducibility. The limits of agreement are shown by the average difference ± 1.96 Std.

samples from these patients. All samples in these experiments are pretreatment.

In Experiment 2 (day 0), 396 proteins were identified and quantified and in Experiment 3 (day 7), 374 proteins. There were 493 unique proteins altogether and 277 of them were present across both data sets. In total across all three iTRAQ experiments, 576 unique proteins were identified and quantified, of which 244 were present in all experiments. The iTRAQ protein ratios for the replicate pooled reference samples in both Experiments 2 and 3 were compared to assess experimental reproducibility across two separate runs of our workflow. Bland-Altman test for the agreement between these experimental replicate pools showed good agreement and therefore no significant differences were found using Pitman's test for differences in variance ($n = 268$, $p = 0.085$, $r = -0.106$) (Figure 2). Therefore the method allowed for the direct comparison of protein ratios across multiple iTRAQ experiments via pooled reference samples, such as would be required in any longitudinal clinical study. It also indicates that technical variance present in the experiments, which is essential for carrying out the *post hoc* power analysis, can be approximated as the average of technical variance present in each individual iTRAQ experiment.

We investigated the fold changes of proteins quantified in both PACER experiments to identify proteins that may be differentially expressed over a 7 day period pretreatment. As the within-person variances of Experiment 2 and 3 were not available, the variance we derived in Experiment 1 was applied in the analysis of Experiment 2 and 3. A protein is considered differentially expressed if its variation within the technical replicates is smaller than the 95% CI defined by the technical variance, whereas its changes of expression level over 7 days period is larger than the 95% CI defined by the technical and within person variance. In total, 29 proteins showed significant

changes in at least one patient (Table 4). It was apparent that Patient D had considerably more differentially expressed proteins than the other two patients, although clinical data for the three patients over the 7 day period does not indicate any obvious confounding factors that may have led to the large changes.

The largest change observed was in Peroxiredoxin II, where a 14-fold increase was observed after 7 days in Patient D, and a smaller yet also significant increase was observed in Patient E. According to the record from Universal Protein Resource (UniProt, <http://www.uniprot.org/>), this protein may be involved in signaling cascades of growth factors and tumor necrosis factor-alpha and is relevant to antiapoptotic processes. Western blotting for Peroxiredoxin II confirmed this protein to be changing, using Coagulation Factor XIII B Chain Precursor as a loading control as this was found to be unchanged across all 3 patients at both time-points in the proteomic analysis (Figure 3).

According to the observed variance, none of the proteins listed in Table 4 changed significantly for all patients (2 sided t test, data not shown). Those showing highest significance and power however included: Ig alpha-1 chain C region (IGHA1), Receptor-type tyrosine-protein phosphatase gamma (PTPRG) and Endoplasmic reticulum chaperone (HSP90B1). This gives an example of the approach that can be used with 8 channel isobaric tagging for clinical proteomics associated with underpinning clinically relevant power analysis. We stress no novel biomarker is immediately apparent from this study, as expected, but Hsp90B1 is a member of the hsp90 family of molecular chaperones, whose inhibition by geldanamycin-derived compounds can activate the unfolded protein response and led to cell death in melanoma cells, exposing a potential route to novel anticancer treatments.

Table 4. Proteins with Differential Expression in the PACER Study between Pretreatment Day 0 (Experiment 2) and Day 7 (Experiment 3)^a

protein names	patient C, day 7:0	patient D, day 7:0	patient E, day 7:0
Anti-(ED-B) scFV (Fragment)	1.395	0.081	0.663
ALDH1A1 Retinal dehydrogenase 1	1.099	3.420	1.469
APOL1 Isoform 2 of Apolipoprotein L1	1.059	0.347	0.901
CA1 Carbonic anhydrase 1	1.542	9.579	1.758
CETP Isoform 1 of Cholesteryl ester transfer protein	1.561	2.031	1.081
CFP Properdin	1.206	2.412	1.257
CRP Isoform 1 of C-reactive protein	1.658	0.383	1.312
FETUB Fetuin-B	0.953	0.361	1.034
GAPDH Glycerinaldehyde 3-phosphate dehydrogenase	0.750	2.107	1.207
GOT1 Aspartate aminotransferase, cytoplasmic	1.227	2.637	1.017
HGFAC Hepatocyte growth factor activator	0.802	2.117	1.595
HSP90B1 Endoplasmic	0.592	0.377	0.634
KRT5 Keratin, type II cytoskeletal 5	0.515	0.477	0.696
PARK7 Protein DJ-1	1.572	3.599	1.444
PDLIM1 PDZ and LIM domain protein 1	1.044	0.333	0.520
PFN1 Profilin-1	0.947	2.771	1.785
PRDX6 Peroxiredoxin-6	1.412	5.408	1.492
PTPRG Isoform 1 of Receptor-type tyrosine-protein phosphatase gamma	1.829	2.557	1.351
SAA2 Serum amyloid A protein	1.151	0.312	0.764
TALDO1 Transaldolase	1.082	3.046	1.154
TPI1 Isoform 1 of Triosephosphate isomerase	0.906	2.137	1.738
TMSB4X TMSB4X protein (Fragment)	0.378	0.382	0.797
PRDX2 Peroxiredoxin-2	1.643	14.362	5.921
CAT Catalase	0.947	8.028	2.132
HBA1 Hemoglobin subunit alpha	1.740	10.522	2.904
HBB Hemoglobin subunit beta	1.580	8.881	3.275
HBD Hemoglobin subunit delta	1.579	8.691	2.429
IGHA1 cDNA FLJ90170 fis	1.678	1.937	2.062
PDLIM7 Isoform 1 of PDZ and LIM domain protein 7	1.516	0.954	2.204

^aProteins identified to be differentially expressed in at least one patient are listed in the table and the significant changes are indicated in bold italic.

The number of patients that are required to reach 70% power were listed in Supplementary Table 4 (Supporting Information). Clearly, most proteins have very high variance (>85 percentile), a feature which is primarily due to between person variation and as such these can hardly be valid candidates for biomarkers (see Discussion section for more detail). For the proteins with less variance, the number of patients required to reach a 70% power estimated using *post hoc* and *a priori* power analysis were compared as illustrated in Supplementary Figure 4. Considerable agreement can be seen between the two methods, confirming the validity of the *a priori* power analysis.

DISCUSSION

There is a clear clinical need for novel predictive, prognostic and/or pharmacodynamic biomarkers in easily sourced material such as plasma. The MS-based method that we have described provides a robust platform to compare multiple proteins simultaneously, to allow identification of novel biomarkers with clinical utility. By use of iTRAQ tagging in conjunction with

extensive statistical testing during data analysis, we have validated a workflow that is applicable to large scale longitudinal clinical trials. Furthermore, the experimental design which includes the use of a pooled reference sample run in duplicate for each iTRAQ experiment clearly demonstrates the utility of this methodology to compare fold changes in protein expression across multiple experiments.

Ernault et al.²⁷ employed an iTRAQ methodology in parallel workflows utilizing immunodepletion or hexapeptide ligand library enrichment to identify 243 and 228 proteins with at least 2 peptides giving a combined total of 313 proteins. The inclusion of single peptide protein identifications would have increased these numbers to 332 and 320 for the immunodepleted and hexapeptide enrichment methods employed, respectively. Kolla et al.²⁸ have used 4-plex isobaric tagging to analyze maternal plasma in Down's Syndrome pregnancies, identifying 187 proteins. Pernemalm et al.²⁹ identified and quantified 300 proteins in an adenocarcinoma plasma study and 193 proteins in a pancreatic cancer study. Thus our approach, which identified 576 unique proteins in 3 iTRAQ runs, is statistically rigorous, yields more protein identifications and has the benefit of 8 samples being analyzed per run. Our data showed MS variance to be low and comparable to that reported elsewhere.^{30,31} The variation levels we reported showed the number of intricate steps involved in our experimental workflow to be robust and reproducible. A study on all potential cancer biomarkers found in the literature showed 49% were present at <10 ng/mL in plasma.³² Therefore our identification of IL6-R, which is present at subng/mL amounts (picomolar levels) in plasma, indicates that our discovery approach has the capacity to uncover potential biomarkers, especially in the context of studies on patients undergoing clinical intervention with pre- and post-treatment samples collected longitudinally. We were also able to confirm changes in Peroxiredoxin II by Western blotting, showing that this protein was up-regulated in Patient D and E (Figure 3). This further validates our workflow design and gives confidence that we can identify novel biomarkers of predictive, prognostic or other clinical use. In addition to this, members of the Peroxiredoxin family (including II) have been linked to pancreatic^{33,34} and other cancers.^{35,36}

It has been suggested that for the determination of reliable identification and quantification of a protein by MS it is necessary for at least two peptides to be identified. However, it is recognized that this may result in the loss of potentially interesting small or low abundance proteins. The inclusion of single peptide data has been debated³⁷ and it has been suggested that a two-peptide or more rule should be replaced by peptide identifications based on thresholds derived from a more statistically robust estimation of error rates.³⁸ This supports the use of our stringent *q*-value¹⁴ based statistical approach to determine peptide confidence levels in order to minimize the number of false positive identifications. By extending our FDR calculations to provide each PSM with its own measure of significance, while accounting for multiple testing this approach provides a robust assessment of the proportion of significant PSMs that turn out to be false positives. This enables the inclusion of single peptide protein identifications and thus maximizes the potential to identify novel low abundance biomarkers for clinical utility.

In this study, the blood proteomes from the pancreatic cancer patients analyzed showed obvious differences. Among the 29 proteins that are differentially expressed in at least one

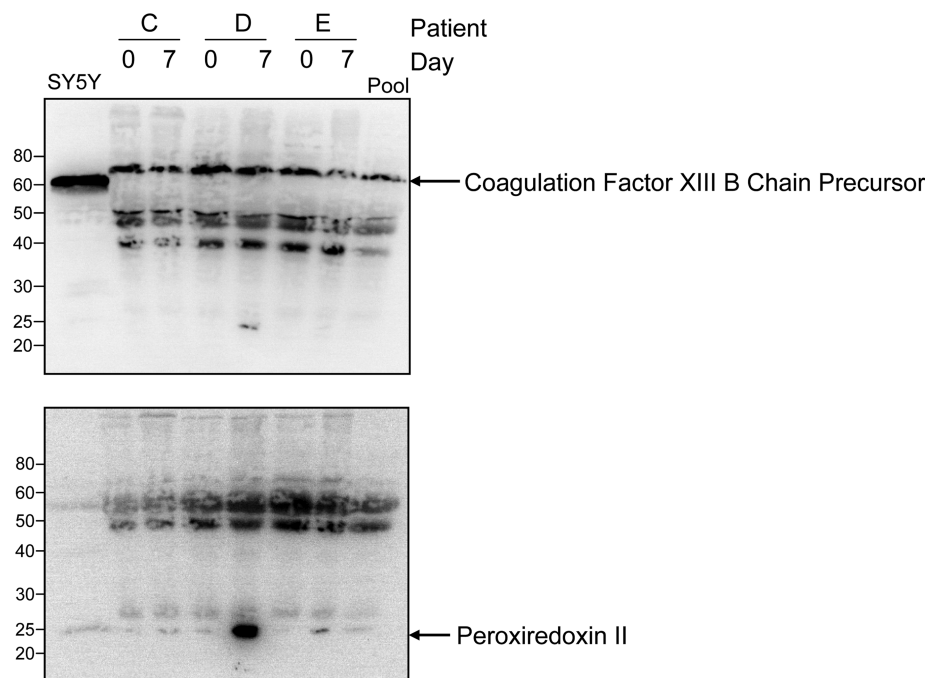


Figure 3. Uncropped Western blots for levels of Peroxiredoxin II and Coagulation Factor XIII B Chain Precursor in undepleted patient plasma. Protein levels are shown in relation to the pooled reference and SH-SY5Y lysates were used as a positive control.

patient, 27 were found changing significantly in Patient D, whereas only 7 were found in Patient E and none in Patient C. No protein changes significantly in all three patients. This is, however, not a surprise because all samples used in Experiment 2 and 3 are pretreatment and there is no indication of clinical difference, such as disease progression, during this period. Thus the proteins that were found differentially expressed are more likely to reflect the clinical condition of each individual rather than act as biomarkers for pancreatic cancer. Although we did observe proteins that may be directly relevant to cancer (prognosis, treatment or response), such as HSP90B1, which is worth further investigation in future studies including post treatment patient samples, this study has highlighted the absolute requirement for measurement of baseline changes in the plasma proteome of patients prior to treatment to distinguish true treatment-related effects.

Typically, iTRAQ proteomics data sets as well as other -omics data sets require high costs (money, time, etc) to produce, especially in experiments involving clinical samples from patients. It is essential to find out the minimum number of patients required to provide enough findings. In this study, we aimed just to clarify the use of power analysis in the context of complex isobaric tagging or relative quantification mass spectrometry, and the purpose of power analysis was expressed as one would find in a well designed clinical study: with the expected fold changes, to determine the number of patients required for a proportion of least variant proteins to have sufficient statistical power to make the study give insight. In a typical biomarker discovery experiment, variant proteins are inherently less likely to be valid candidates for a universally applied biomarker. We can propose that 6 patients per cohort, allowing for 2 fold changes plus 70% power for 80% of the least variant proteins, will be a sufficient starting point for a robust biomarker discovery experiment. It asks for experimental capacity that is entirely tractable with the current technology, and maintains a reasonable level of expected statistical power.

Following candidate biomarkers identified by this method, targeted investigations may be carried out on additional patient samples that may also be required in order to verify proteins with higher variance or to obtain greater statistical power.

We stress that the number of patients required to get sufficient statistical power has been calculated by both *a priori* and *post hoc* power analysis. Comparison showed considerable agreement between the two results (Supplementary Figure 4, Supporting Information) for proteins with lesser variance (<85th percentile), which are of primary interest in biomarker studies. Essentially, such agreement indicates that despite differences in experimental condition, disease type, etc., the variance range for the majority of proteins does not vary significantly. Therefore the results from the *a priori* power analysis (Table 3) can be applied universally for future iTRAQ experiments.

In this paper, we have described a framework by which clinical proteomic study designs can minimize the FDR in protein identification and quantification, leading to thorough statistical assessment of technical and biological variation on a study by study basis. This replaces the use of arbitrary thresholds based upon variance levels reported in other studies which may be completely unrelated. It is critical to provide a robust assessment of both technical and biological variance, and in doing so here we have highlighted the importance of accounting for these errors during data analysis. Thus, we have validated the methodology for clinical trial proteomics and provide a power analysis solution which falls realistically into study design parameters for clinical trials.

■ ASSOCIATED CONTENT

📄 Supporting Information

Supplemental tables and figures. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*Prof. Caroline Dive, Tel: +44 (0)161 446 3036. Fax: +44 (0) 161 446 4109. E-mail: CDive@picr.man.ac.uk Prof. Anthony D. Whetton, Tel: +44 (0)161 275 0039. E-mail: tony.whetton@manchester.ac.uk

Author Contributions

§These authors contributed equally to this work

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Dr. Kathryn Lilley and Dr. Andrew Williamson for their valuable comments on the manuscript. C.D., C.Z., L.J.L., K.L.S. and M.J.D. were funded by Cancer Research UK (Paterson Institute for Cancer Research Core Funding from Grant Code: C147/A12328, Clinical Research Initiative Grant Code: C357/A12197); PACER funding was from Feasibility Study Committee (C153/A7727). PACER-TRANS funding was from the Experimental Cancer Medicine Centre Network (ECMC) to A.R. R.D.U. is funded by NIHR Manchester Biomedical Research Centre, A.D.W. and R.D.U. are funded by Leukaemia Lymphoma Research UK (LRF code 08004) and Cancer Research UK and M.W. is funded by the ECMC.

REFERENCES

(1) Beretta, L. Proteomics from the clinical perspective: many hopes and much debate. *Nat. Methods* **2007**, *4*, 785–6.

(2) Ross, P. L.; Huang, Y. N.; Marchese, J. N.; et al. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* **2004**, *3*, 1154–69.

(3) Pierce, A.; Unwin, R. D.; Evans, C. A.; et al. Eight-channel iTRAQ enables comparison of the activity of six leukemogenic tyrosine kinases. *Mol. Cell. Proteomics* **2008**, *7*, 853–63.

(4) Wu, W. W.; Wang, G.; Baek, S. J.; Shen, R. F. Comparative study of three proteomic quantitative methods, DIGE, cICAT, and iTRAQ, using 2D gel- or LC–MALDI TOF/TOF. *J. Proteome Res.* **2006**, *5*, 651–8.

(5) Karp, N. A.; McCormick, P. S.; Russell, M. R.; Lilley, K. S. Experimental and statistical considerations to avoid false conclusions in proteomics studies using differential in-gel electrophoresis. *Mol. Cell. Proteomics* **2007**, *6*, 1354–64.

(6) Abdi, F.; Quinn, J. F.; Jankovic, J.; et al. Detection of biomarkers with a multiplex quantitative proteomic platform in cerebrospinal fluid of patients with neurodegenerative disorders. *J. Alzheimers Dis.* **2006**, *9*, 293–348.

(7) DeSouza, L. V.; Grigull, J.; Ghanny, S.; et al. Endometrial carcinoma biomarker discovery and verification using differentially tagged clinical samples with multidimensional liquid chromatography and tandem mass spectrometry. *Mol. Cell. Proteomics* **2007**, *6*, 1170–82.

(8) Hergenroeder, G.; Redell, J. B.; Moore, A. N.; et al. Identification of serum biomarkers in brain-injured adults: potential for predicting elevated intracranial pressure. *J. Neurotrauma* **2008**, *25*, 79–93.

(9) Kolla, V.; Jenö, P.; Moes, S.; et al. Quantitative proteomics analysis of maternal plasma in Down syndrome pregnancies using isobaric tagging reagent (iTRAQ). *J. Biomed. Biotechnol.* **2010**, 952047.

(10) Ogata, Y.; Charlesworth, M. C.; Higgins, L.; Keegan, B. M.; Vernino, S.; Muddiman, D. C. Differential protein expression in male and female human lumbar cerebrospinal fluid using iTRAQ reagents after abundant protein depletion. *Proteomics* **2007**, *7*, 3726–34.

(11) Unwin, R. D.; Smith, D. L.; Blinco, D.; et al. Quantitative proteomics reveals posttranslational control as a regulatory factor in primary hematopoietic stem cells. *Blood* **2006**, *107*, 4687–94.

(12) Williamson, A. J.; Smith, D. L.; Blinco, D.; et al. Quantitative proteomics analysis demonstrates post-transcriptional regulation of embryonic stem cell differentiation to hematopoiesis. *Mol. Cell. Proteomics* **2008**, *7*, 459–72.

(13) Elias, J. E.; Gygi, S. P. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat. Methods* **2007**, *4*, 207–14.

(14) Storey, J. D.; Tibshirani, R. Statistical significance for genome-wide studies. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 9440–5.

(15) Käll, L.; Storey, J. D.; MacCoss, M. J.; Noble, W. S. Assigning significance to peptides identified by tandem mass spectrometry using decoy databases. *J. Proteome Res.* **2008**, *7*, 29–34.

(16) Daly, D. S.; Anderson, K. K.; Panisko, E. A.; et al. Mixed-effects statistical model for comparative LC–MS proteomics studies. *J. Proteome Res.* **2008**, *7*, 1209–17.

(17) Demirkale, C. Y.; Nettleton, D.; Maiti, T. Linear Mixed Model Selection for False Discovery Rate Control in Microarray Data Analysis. *Biometrics* **2009**.

(18) Munro, N. P.; Cairns, D. A.; Clarke, P.; et al. Urinary biomarker profiling in transitional cell carcinoma. *Int. J. Cancer* **2006**, *119*, 2642–50.

(19) Karp, N. A.; Lilley, K. S. Maximising sensitivity for detecting changes in protein expression: experimental design using minimal CyDyes. *Proteomics* **2005**, *5*, 3105–15.

(20) Dobbin, K.; Simon, R. Sample size determination in microarray experiments for class comparison and prognostic classification. *Biostatistics* **2005**, *6*, 27–38.

(21) Storey, J. D.; Tibshirani, R. Statistical significance for genome-wide studies. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 9440–5.

(22) Alexandrakis, M. G.; Passam, F. H.; Boula, A.; et al. Relationship between circulating serum soluble interleukin-6 receptor and the angiogenic cytokines basic fibroblast growth factor and vascular endothelial growth factor in multiple myeloma. *Ann. Hematol.* **2003**, *82*, 19–23.

(23) Rifai, N.; Gillette, M. A.; Carr, S. A. Protein biomarker discovery and validation: the long and uncertain path to clinical utility. *Nat. Biotechnol.* **2006**, *24*, 971–83.

(24) Breitwieser, F. P.; Muller, A.; Dayon, L.; et al. General statistical modeling of data from protein relative expression isobaric tags. *J. Proteome Res.* **2011**, *10*, 2758–66.

(25) Gan, C. S.; Chong, P. K.; Pham, T. K. Wright PC. Technical, experimental, and biological variations in isobaric tags for relative and absolute quantitation (iTRAQ). *J. Proteome Res.* **2007**, *6*, 821–7.

(26) Yang, Y. H.; Speed, T. Design issues for cDNA microarray experiments. *Nat. Rev. Genet.* **2002**, *3*, 579–88.

(27) Ernout, E.; Bourreau, A.; Gamelin, E.; Guette, C. A proteomic approach for plasma biomarker discovery with iTRAQ labelling and OFFGEL fractionation. *J. Biomed. Biotechnol.* **2010**, 927917.

(28) Kolla, V.; Jenö, P.; Moes, S.; et al. Quantitative proteomics analysis of maternal plasma in Down syndrome pregnancies using isobaric tagging reagent (iTRAQ). *J. Biomed. Biotechnol.* **2010**, 952047.

(29) Tonack, S.; MA-O, D.; Jenkins, R. E.; et al. A technically detailed and pragmatic protocol for quantitative serum proteomics using iTRAQ. *J. Proteomics* **2009**.

(30) Song, X.; Bandow, J.; Sherman, J.; et al. iTRAQ experimental design for plasma biomarker discovery. *J. Proteome Res.* **2008**, *7*, 2952–8.

(31) Gan, C. S.; Chong, P. K.; Pham, T. K. Wright PC. Technical, experimental, and biological variations in isobaric tags for relative and absolute quantitation (iTRAQ). *J. Proteome Res.* **2007**, *6*, 821–7.

(32) Polanski, M.; Anderson, N. L. A list of candidate cancer biomarkers for targeted proteomics. *Biomark. Insights* **2007**, *1*, 1–48.

(33) Cecconi, D.; Donadelli, M.; Rinalducci, S.; et al. Proteomic analysis of pancreatic endocrine tumor cell lines treated with the histone deacetylase inhibitor trichostatin A. *Proteomics* **2007**, *7*, 1644–53.

- (34) Park, J. Y.; Kim, S. A.; Chung, J. W.; et al. Proteomic analysis of pancreatic juice for the identification of biomarkers of pancreatic cancer. *J. Cancer Res. Clin. Oncol.* **2011**, *137*, 1229–38.
- (35) Basu, A.; Banerjee, H.; Rojas, H.; et al. Differential expression of peroxiredoxins in prostate cancer: consistent upregulation of PRDX3 and PRDX4. *Prostate* **2011**, *71*, 755–65.
- (36) Woolston, C. M.; Storr, S. J.; Ellis, I. O.; Morgan, D. A.; Martin, S. G. Expression of thioredoxin system and related peroxiredoxin proteins is associated with clinical outcome in radiotherapy treated early stage breast cancer. *Radiother. Oncol.* **2011**, *100*, 308–13.
- (37) Bradshaw, R. A.; Burlingame, A. L.; Carr, S.; Aebersold, R. Reporting protein identification data: the next generation of guidelines. *Mol. Cell. Proteomics* **2006**, *5*, 787–8.
- (38) Gupta, N.; Pevzner, P. A. False discovery rates of protein identifications: a strike against the two-peptide rule. *J. Proteome Res.* **2009**, *8*, 4173–81.