Global Biomarkers Standardization Consortium of CSF biomarkers

AAIC Face-to-Face Meeting

July 15, 2017

London, UK

Meeting Summary

Welcome and Introduction

- Jim Hendrix gave the welcome and extended accolades to the Consortium for being the driving force in CSF research and for this year's, best-attended, Biofluids Workshop.

Update on QC Program - Kaj Blennow

- Initiated Alzheimer's Association CSF program with Alzheimer's Association funding, the program was started due to the large variability of absolute levels of CSF Aβ42 across labs displaying the need for standardization efforts.
- Variability is due to pre-analytical factors, analytical factors, and assay manufacturing.
- Started in 2009, 3 rounds per year, 3 QC samples (pooled CSF; 2 unique, 1 same).
- Goals for QC program are for individual labs to monitor: accuracy of CSF biomarker measurements; longitudinal drift in CFS levels; assay performance and batch variation.
- 24 rounds completed to date. 126 total number of labs, numerous assay formats: Innotest ELISA, Luminex, MesoScale bead-based assay, Euroimmun ELISA, Cobas Elecsys, IBL (too few participants for evaluation), Lumipulse from Fujirebio.
- Testing of CSF Aβ(1-42) on Cobas Elecys
    - Minimal batch to batch variability; method aligned to reference measurement procedure of mass spectrometry.
    - Same round comparisons of in between-lab variability: Innotest-17% CV and Cobas Elecys-2.5% CV.
- Testing of CSF Aβ42 on Lumipulse G1200 and Lumipulse G600 II
    - Method aligned to reference measurement procedure of mass spectrometry
    - First round gave a marked reduction in between-lab variability.
- An example was given monitoring between-batch variability for the Innotest ELISA- Round 24, 5 different batches, no differences in Aβ(1-42), T-tau and P-tau.
- Results from last rounds 2016-2017

Assay	CSF Aβ42 Mean CV (%)	CSF Aβ40 Mean CV (%)	CSF T-tau Mean CV (%)	CSF P-tau Mean CV (%)
Innotest ELISA	15	27	15	12
Euroimmun ELISA	13	16	11	
AlzBio3 Luminex	22		14	40
Meso-Scale ECL V-Plex	20	16		
Elecys			4	2

***gray box denotes not reported in presentation*

- The QC program will continue to monitor the performance of the AD CSF biomarker

assays between lab CVs and longitudinal changes.

- When the CSF reference material is available will Kaj be involved in facilitating the QC program using the material? Kaj answered the aim of the reference material is not to place it in the QC program; it is when you are trying to standardize your assays during production or with a new batch, etc.
- Hugo asked if Kaj can put a level of acceptance for assays in such a program. With the CVs, what is still acceptable? Must be consensus agreement either with the GBSC or clinicians on what is acceptable. Should not be too strict, for ex. that anything above 1% is unacceptable.
    - Charlotte utilizes it for quality purposes, the limits get predefined by themselves, the goal is 15% within the limits, 1 SD of the average of the total Innotest performance.
- Ingrid reported that the European Federation for Laboratory Medicine published guidance on what the criteria should be of acceptance of a method such as based on non-clinical impact on variation of the method, biological variation, etc.
- Once certified reference material is released and manufacturers can re-standardize their assays, will Kaj start working on target values? Kaj will have to discuss this.
- Bob commented in regards to setting required expectations, there are country specific requirements set for clinical labs based on analytic methodology and context of expected biological variability. To get to a universal cut point/reference interval there is still a lot of work that needs to be done on pre-analytics; it is premature to set hard expectations for performance. As a group, strive to influence clinical research; try to avoid mandated analytical performance criteria because in many cases this has implications in lab reimbursement.
- A member recommended to first start on stable assays, repeat a study measuring analytes that had low variations within subjects at 6 months, repeat this study with one of the robust assays and re-asses inter-individual variation. Also, if any drugs come out with benefits, such as detection of reduction of tau, they can tell how large the reduction is and can use % reduction to learn how precise methods have to be.
- Majority of the clinical testing facilities in Germany that are tested in external QC schemes and legally binding, in regards to acceptance criteria, have hard cut-offs, if not fulfilled will have to stop analyses and will not be reimbursed.

Reference Material: IRMM/IFCC Project Overview - Henrik Zetterberg & Ingrid Zegers

- 3 candidate serums, processed and produced and characterized in terms of homogeneity, each batch has vials with the same concentration of Aβ42. Have assessed short-term stability for shipment and long-term stability for storage at -70°C.
- LC-MS Ring Trial II- characterization of the reference material using reference measurement procedures from Ring Trial. UGot, Roche, UPenn, Waters and PPD companies used common Aβ42 calibrant but different protocols for preparing calibration solutions, measurements over 3 days, 3 samples/day, either in triplicate or duplicate.
- Results: Graphs displayed no values because material is not formally released but labs reported consistent results. The material will be certified. CV between labs is between 4-7%; between lab variation and between-day variation are main contributions of uncertainty.
- Reference method is under control as could be for the analyte.
- The 3 reference materials are commutability for a series of methods.
- Prepared documentation currently being assessed by external experts, next steps: finalize external review and figure out how to use the materials.
- Kaj inquired if Ingrid plans to distribute a calibrator of Aβ peptide? This has been a discussion point within the IFCC working group; the labs that run the reference measurement procedure are interested in having a calibrator, but not other labs.
- Henrik provided an update of 3 new projects: Candidate reference method for Aβ40, CRM for tau, pilot commutability study for tau
- Amyloid 1-40 Candidate RMP

    - MS based method, endogenous Aβ40 and an isotope internal standard added to quantify Aβ40.
    - Results show a linearity of the method over a wide range of Aβ40 concentrations.
    - Precision is good with low error and sample is stable with freeze/thawing.
    - Method submitted for evaluation to the Joint Committee for Traceability in Laboratory Medicine.
- CRM for tau
    - Selected a reporter fragment that has no post-translation modifications and no risk of oxidation-should be a stable reporter molecule.
    - Digested CSF results-Digest with trypsin and measure reporter fragment. Graph indicates that digested tau was being measured, not a contaminant or endogenous peptide.
    - Native tau in CSF- t-tau ~650 pg/ml, CV-4.6%.
    - Innotest ELISA-measured 3 concentrations of t-tau, MS signals corresponded with concentration levels.
    - Next steps are: Determine if full length proteins should be used for calibration or determine actual concentrations of aliquots using amino acid analysis; skip protein precipitation; validation of the method.
- Pilot commutability for tau
    - 3 candidate CRMs for Aβ42 were assayed for T-tau, 34 individual CSF samples were also included, samples analyzed by the labs that developed the respective tau method.
    - Nice Correlations between IBL vs. Roche assays, Euroimmun vs. Roche, and MesoScale vs. Roche, Innotest vs. Roche, Euroimmun vs. IBL, and MesoScale vs. IBL.
    - Conclusions: CSF Aβ42 project almost completed; promising data on CRM for Aβ40 and tau; promising commutability.
    - May not need to have human CSF for tau; spiked artificial CSF may work, a formal commutability study of different CRMs should be performed.

Integration of reference materials into production flow of assays - Britta Brix

- After the release of the CRM what will happen once the vendor receives it?
- From IRM to re-calibrated assays- the approach for Aβ42
    - The system consists of the CRM, assay and the process
- Will receive 3 CSF samples; high, medium and low. One method would be to dilute sample to get a calibrator series, also get an internal CSF series, use reference series on internal CSF series to obtain the internal calibrators that will put new values to the samples. There are alternative methods as well and all the different processes can have an influence on the final value.
- Recommended a working group for re-calibration, streamline the re-calibration, run round-robin before market entry, enter market with new values and preanalytical SOP
- Jim Hendrix asked the members if they are interested in joining the working group to notify him.
- Ingrid informed the group that there will not be instructions given on how to use the CRM.

Biofluid Based Biomarkers PIA - Sid O'Bryant & Henrik Zetterberg

- Expanded PIA focus from solely blood to biofluid based biomarkers
- Short-term goals include working towards guidelines, best practices, harmonization, etc.
- Long-term goals have been expanded to trial based appropriate outcomes.
- PIA is working to learn largely from the GBSC; with the expansion now have several GBSC members on the Executive Committee.
- Currently looking for someone who has expertise in either saliva or urine based biomarkers: if interested or know of someone who does have that expertise, contact Sid.
- Published guidelines for preanalytic processing focused on serum and plasma.
- Group created a paradigm on how to move blood based biomarkers forward.

- 3rd largest PIA, ~600+members
- 4 projects have come forward based on the PIA Day meeting: harmonization and standardization of preanalytics factors in saliva; looking at analytics, how are the bioinformatics methods being used, what is appropriate; context of use and what are the context of use of greatest need; looking at biomarkers from an epidemiological standpoint.
- Sid notified the group that if anyone is interested in any of those projects or taking a lead, let the Executive Committee know.

Alzheimer's Disease Center CSF Standardization Efforts - Nina Silverberg

- Director of Alzheimer's Centers Program, NIA.
- Recently completed a year-long process developing recommendations for the future of the Alzheimer's Centers.
- 30 Alzheimer's Centers across the country each see 100s of participants annually and collect data that is sent to the National Alzheimer's Coordinating Center in Seattle. A proportion of participants have an autopsy performed. A lot of samples available, such as imaging and biomarkers, variable across Centers. Data from diverse populations and other dementias.
- Currently, difficult to find out what each center has and what might be available for which purposes. Nina wants to develop a workgroup to assist with this and wanted to inform the GSBC of this opportunity.

LP Video - Charlotte Teunissen

- Society for CSF Analysis and Clinical Neurochemistry Symposium-June 7-8, 2018.
- Published a LP video in Alzheimer's and Dementia Diagnosis Assessment and Disease Monitoring.
- Current video is for professionals and the next step is a video for patients to reduce the risk of post puncture complications.
- Significant risk factors for PLPH-ranked by magnitude-4000 subjects, with both patient related and procedure related factors, highest magnitude for each is history of severe/chronic headache and cutting needle type but the patient being very worried is also a risk factor.
- Very worried was a top ranked risk factor for non-specific headache and slightly worried was listed. Also, being worried had an effect on back pain.
- Animation video is recommended because it can be easily relatable to older age, gender, different languages, etc.
- John suggested changing terminology from lumbar puncture to CSF collection (has been adapted in Washington University) patients are more accepting to it.
    - The strongest feedback from the DIAN meeting was that LP terminology is not good.
- Video was well received but different countries have different practices, ex. the use of anesthesia.
- UK Alzheimer's Society has a patient video on YouTube and another one of a younger patient with Huntington's disease undergoing LP.

Automated Platforms - Ulf Andreasson

- Different levels of automation: no automation-manual pipetting and preparation of reagents; semi-automated-robotic pipetting but manual preparation of reagents; fully automated-robotic pipetting and ready-to-use reagents.
- Semi-automated examples-Tecan, Hamilton, liquid handling robot, washer, etc.
- Fully automated
    - RA Analyzer 15 (Euroimmun)-Coming out soon, not yet in QC program. 101 tests/hr. result in less than 30 min., chemiluminescence detection**.**
    - Lumipulse G1200 and Lumipulse G600 (Fujirebio) - 120 tests/hr. result in less than 30 min., chemiluminescence detection.
    - Cobas e 601 (Roche)-results in less than 20', 120test/hr,electrochemiluminescence

detection.

- Release of AD markers on fully automated platforms:
    - Euroimmune - Aβ42- Q4 2017, Aβ40- Q4 2017, T-tau and P-tau-expected soon
    - Roche - Aβ42- Released, Aβ40-prototype, T-tau and P-tau-expected soon
    - Fujirebio - Aβ42- Released, Aβ40-prototype, T-tau-will be released tomorrow, July 16, 2017, and P-tau-prototype
- Discussion was if the future direction is that samples will be sent to a central lab for testing.

Cut-points for AD biomarkers: Methods & Challenges - Jonathan Schott

- A binary test is preferred; you do or do not have the disease.
- In research an increasing move to use biomarker designations-A/N scheme, A/T/N scheme. There is a need for a cut-point.
- A publication displayed different cut-points for different European centers, a patient would be diagnosed with Alzheimer's differently in different countries due to the variable cut-points
- Few tests are binary, exception genetics, but even in genetics there is a gray zone
- Methods for determining cut-points depend on, or change, sensitivity/specificity balance:
    - Can establish normal reference limits based just on controls. Can maximize accuracy, depends on disease prevalence. Receiver operator curve methods-Youden's index, mixture modeling-all easy when 2 separate Gaussian curves, difficult when overlap.
- For a control group to actually be a control, need another biomarker because 20-30% of elderly individuals may have prodromal AD.
- If you use younger controls, assume that your measures do not change "normally" with age.
- Who is the patient group? Pure AD is rare many have more than one pathology. In clinical practice the issue is not distinguishing AD dementia from controls, but AD from other forms of dementia.
- Beach et al 2012 - postmortem samples, 107/271 diagnosed with non-AD had AD.
- AD have different presentations, pathology, genetic influences and may have different biomarker characteristics.
- Best thing to do is use post-mortem confirmed samples; example is the ADNI data set.
    - Tau is tuned for specificity and not sensitivity (will miss cases) Aβ is very sensitive and not very specific. If apply A/T/N scheme using those cut-points will get very different measures for each in regards to who is positive or negative.
- Clinically derived methods: assessment of disease progression-takes time, good for determining prognosis/progression. Comparisons with other biomarkers- question as to which is the gold standard?
- Measured Aβ42, Aβ42/40, tau/Aβ42 and p-tau from people with a quantitated amyloid scan and a clinical dichotomized rating (positive or negative) and looked at concordance.
- Discordant-gray zones, need to look at gray zones. Not a definite cut-point, has to be a gradation.
- Cut-points have changed in blood pressure goals and diabetes.
- Cut-points are important but imperfect, need to take this in to account in new criteria for AD.
- Sensitivities and specificities for different levels for different stages.
- Define labels better- ex. amyloid PET
- With Aβ, a negative predictive value may have more certainty.
- Need larger biomarker comparison studies, pools of biomarker healthy controls, and long term follow up, and obtain as much autopsy information as possible.
- Les Shaw inquired on what is meant by biomarker normal individuals? If we obtain CSF from amyloid negative controls, understand may have age related tauopathies, etc., it would be useful.

Panel Discussion Ulf Andreasson and Jonathan Schott

- Les Shaw-In regards to cut-points thoughts about statistical approaches? Jonathan stated this is rather difficult, amyloid imaging provides nice separation; concern is comparing different biomarkers together when using different sensitivities and specificities, particularly when trying to put them in to a common framework.

- Henrik Zetterberg stated that one approach is to work with likelihood ratios but need to standardize assays.
- Different pre-analytics cause a problem, different cut offs in different populations
- Jonathan suggested for the group to work towards a central online resource that allows the user to input a value and receive a probability of a diagnosis; the probability would be weighted more when more post-mortem samples are included. This would become more accurate based on the more data inputted.
- Bob-need to be very clear on the question trying to answer. Ex. prognostic questions, do I have someone with a chemical finding that will predict a neuroimaging finding, or pathology finding? Will probably have to lower the expectations on biomarkers, need to use multiple data.
- Jim Hendrix inquired if there is a need for a similar descriptive label for CSF like Amyvid? Jonathan thinks this is a good starting point.
- In Europe, very country specific, CSF sampling tests are used in clinical practice. Clinical utility in Scandinavia, Germany, France, is already there, not stand alone, adjunct measures.
- How to use 3 and 4 biomarkers together? Need to integrate data with clinical and imaging formation.
- In Germany, biomarkers integrated in clinical routine for at least 15 years, performance of biomarkers is fully reimbursed by insurance. Gray zones have been integrated in to their interpretation algorithms for at least 10 years. ~30% of patients diagnosed with AD had LP.
- Which technique should be used to estimate cut-point? Les and several European research groups used as many methods and narrow in on a reasonable cut-point. Statistics will give you a sliding scale between sensitivity/specificity. Decision is do you want to minimize false negatives or false positives or go for accuracy, not a problem as long as it is acknowledged.
- Might need a staged approach for sensitivity and specificity.
- How to get LP more popular in USA? In France, you have to demonstrate that is saves money. In Britain, need champion physicians to demonstrate ease of use, in UK there is a training program for nurses, and need facilities for practicality.
- Maria Carillo stated that for the US government-based reimbursement agency, it is illegal to talk about money, only benefit to the patient. The burden is to demonstrate the diagnosis improves the health outcome. She recommended for the group to attend the IDEAS Study presentation by Gil Rabinovici, which will provide preliminary results of amyloid PET, proving that amyloid imaging can provide a benefit to a health care management of a patient after diagnosis. Need a FDA approved assay; AUC that can be given to the reimbursement agency; and education beyond that.

Concluding Remarks

- Next webinar teleconference will be scheduled in the fall.