Forum for Collaborative HIV Research
Liver Forum #2
Summary of Proceedings
Tuesday, 21 April 2015
Vienna, Austria

The discussion summarized in this report is based on the perspectives of the various participants. It does not represent the official position of any organization. Specifically, comments made by FDA and EMA representatives are the opinion of the participant rather than official regulatory guidance.

Meeting presentation slides and other background materials can be found here:
http://www.hivforum.org/projects/drug-development/liver-forum

Session #1: Project Overview and Updates
Moderators: Gary Burgess, Veronica Miller and Arun Sanyal

- Welcome and Introductions:
  - Veronica Miller welcomed all present to the 2nd Liver Forum and outlined the content of the first session
  - Gary Burgess followed with his welcome and noted that with each meeting of the Liver Forum, "we seem to need a bigger room". He also noted that when reviewing the slide presentations for this meeting, he was struck by how much the Liver Forum has accomplished since its last meeting.
  - Veronica acknowledged members of the Steering Committee
  - She then thanked the Liver Forum sponsors, especially including AASLD and EASL, as well as the sponsoring pharmaceutical, diagnostic and biotech companies
  - Veronica encouraged participants, especially those from academic institutions to refer their students to the Forum for visiting scholarships or internships.
  - Finally, she introduced and acknowledged the staff present, including Rob Besaw and Nivedha Paneer, the Liver Forum's new epidemiology consultant, Myrna Cozen, and thanked Margie and David Poole and the IHL team for logistical support.
- Operating Principles (Veronica Miller): please, also see:
  http://www.hivforum.org/storage/documents/2015/Liver_Forum/liver%20forum%20operating%20principles.pdf
  - The purpose of The Liver Forum is to encourage collaboration among all sectors involved in developing diagnostic and therapeutic approaches to metabolic liver disease, including non-alcoholic steatohepatitis (NASH), liver fibrosis and cirrhosis
  - Although conducted in conjunction with EASL or AASLD, the Liver Forum is an independent and neutral venue: It doesn't belong to any one sector; it belongs to everybody
  - The Liver Forum is collectively owned by its members, with each stakeholder group having an equal voice
    - It is open to participation by organizations that have an honest commitment to the field, but closed to journalists, investors and marketers
    - Participation is limited to two representatives from each company for in-person meetings
  - It provides a safe place for discussion, dialogue and deliberation; we want all perspectives to be heard: Chatham House Rules are followed

    - Recommendations developed by the Liver Forum must be based on sound ethical principles and scientific evidence
    - We let consensus evolve; we acknowledge when more data are needed before recommendations can be made
    - The Liver Forum promotes "information democracy", the cross pollination of ideas, standardization of research practices and the more efficient design and management of clinical trials. This takes the form of
        - linking the various silos
        - using meetings and fora as mechanisms to encourage interaction and collaboration
        - encouraging everyone to contribute
        - members participating freely and not officially representing their organization
- Overall Goals and Objectives (Arun Sanyal): The purpose of the Liver Forum is to
    - promote ongoing multi-stakeholder dialogue
    - provide the knowledge base that will help everybody with their development programs—whether in the diagnostic or therapeutic space
    - identify gaps in knowledge and development
    - prevent duplication of effort; identify and eliminate inefficiencies; enhance clarity
- Pragmatic Goals: What we actually want the Liver Forum to achieve in the short term (Gary Burgess):
    - One of our big challenges is to identify suitable biomarkers for NAFLD/NASH
    - Biomarkers are needed
        - For both diagnosis and prognosis
        - For selection of appropriate patient populations for clinical studies and for defining inclusion and exclusion criteria
        - To measure how well intervention is performing with regard to disease progression
    - We need to establish how (the various) biomarkers work across different diagnostic modalities to define disease stage
        - Biomarkers need to be predictive, pharmacodynamic and reliable enough to eventually be used as surrogate endpoints
    - Biomarkers are instrumental to the establishment of a natural history cohort
- Progress Report (Veronica Miller): Progress made since the last Liver Forum meeting in November 2014, includes:
    - Issuance of summary report (available on HIV Forum website)
    - Recruitment of new members and volunteers for Working Groups, including leadership for each, and initiation of Working Group discussions
    - Completion of a membership survey identifying gaps in research and prioritizing areas of concentration for this forum
    - Current status:
        - In initial stages of discussion regarding key issues
        - In process of identifying topics for group manuscripts/publications
        - Established and expanded funding base
        - Recruited new staff
- Objectives for today's meeting (Arun Sanyal):
    - Describe and discuss the regulatory perspective; obtain input and feedback from the European Medicines Agency (EMA), represented by Elmer Schabel and the FDA (represented by Lara Dimick)
        - Review progress made by each of the three Working Groups: Data standardization, Case definitions, and Issues in pediatric NASH/NAFLD
        - In each case, the issue of biomarkers will figure prominently
        - We will also explore opportunities for collaboration and working across sectors as well as among institutions

Session #2: Regulatory Perspectives and Updates
Moderators: Gary Burgess and Arun Sanyal

- Abbreviations used in the following presentations:
  - EMA: European Medicines Agency
  - CHMP: Committee for Medicinal Products for Human Use
  - SAWP: Scientific Advice Working Party
  - MAA: Marketing Authorization Application
  - CDER: FDA's Center for Drug Evaluation and Research
  - SPA: Special Protocol Assessment
  - CDRH: Center for Devices and Radiological Health
  - DGIEP: Division of Gastroenterology and Inborn Errors Products

- The European procedure for qualification for novel methodologies for medicine development, presented by Elmer Schabel (The slide presentation can be found here: http://www.hivforum.org/storage/documents/2015/Liver_Forum/schabel_presentation_liverforum2.pdf). Summary remarks and discussion points follow.
  - The qualification procedure is divided into two distinct parts: qualification opinion and qualification advice:
    - Qualification opinion (issued after results of clinical testing are reviewed and applicant is ready to apply for final qualification): CHMP can issue an opinion on the specific use of a novel methodology or imaging method in the context of research and development
      - Method can be applied to non-clinical or clinical studies
      - Before final adoption of an opinion CHMP makes its evaluation available for public comment by the scientific community
    - Qualification advice (issued in response to submission of protocols and preliminary data): CHMP can issue advice on protocols and methods that are intended to be used in developing a novel method with the aim of qualification. This is confidential and no information is disclosed to the public.
      - Advice is based on the evaluation of the scientific rationale and the quality of data submitted to the Agency
  - Biomarker Qualification Procedure is open not only to companies, but also to learned societies and public/private partnerships
    - Groups, including partners from multiple sectors, can apply
    - Not focused on specific products or applications
    - Applicants submit
      - protocols, full study reports and supportive data for qualification or
      - preliminary data and draft protocols for advice
    - Operations include:
      - Pre-submission meetings regarding the purpose of the procedure
      - EMA appointment of a qualification team, led by a coordinator
      - When a qualification opinion is issued, it is followed by
        - public consultation phase, during which time the applicant and the public comment on final qualification opinion to be published
        - information, preliminary reports, comments and final qualification advice being made available on the EMA website
    - Time Course (refer to corresponding slide, page 5 in PDF of PowerPoint presentation, for detail):
      - A qualification advice usually ends at day 100
      - This can be extended if a letter of support is going to be published
      - A lengthy consultation phase may extend the time to up to one year
  - Assessment of experience with qualification procedure thus far:

- Qualification Opinion: 7 reports have been published
  - Most development has taken place in Alzheimer's Disease, for which 4 reports have been published
  - Other single published reports include: methodologies for renal toxicology studies; a statistical model for dose-finding; proposed use of hollow fiber systems (HFS) for TB drug development; and an outcome measure for COPD
- Letters of Support: 3 biomarkers procedures have been finalized: skeletal muscle injury, drug induced kidney injury, microaneurysm rate
- Qualification advice has been issued for 59 procedures since 2007
  - These include: 42 for pure biomarkers; 9 for clinical safety biomarkers; 6 for pre-clinical toxicity biomarkers
  - In addition: 11 procedures related to clinical scores and endpoints and 7 procedures relating to patient-reported outcomes as endpoints
- Large increase in activity since 2010, with more than 20 procedures started within the last year
- Liver related activity: only two procedures on liver safety with regard to drug-induced liver injury have been initiated

- FDA Perspectives and Updates: Experience with Recent Applications, presented by Lara Dimick-Santos. Lara provided an update on FDA activity over the past few months. An interactive dialogue with participants follows.
  - FDA and EMA rules differ, although the qualification process for a biomarker is quite similar and the two agencies work in conjunction with each other
    - FDA does not require that a biomarker be qualified to use it for accelerated approval
    - Whether a biomarker is adequate or appropriate for accelerated approval is a division-by-division decision
  - There is some confusion about the accelerated approval pathway and how it compares with the regular or full approval process. These pathways were described in the report of the FDA-AASLD workshop (Sanyal AJ, et al., Hepatology, vol. 61, no. 4, 2015). Some additional points of clarification:
    - For regular approval, a clinical benefit (i.e., how a patient feels, functions or survives) endpoint is required
      - In some cases, a surrogate endpoint may be accepted for regular approval, if it *does* predict (versus is reasonably *likely to* predict) a clinical endpoint: e.g., hypertension is accepted for predicting stroke in cardiovascular outcomes and creatinine levels are accepted as a predictive biomarker of renal failure
    - For accelerated approval, a clinical benefit (or its accepted surrogate) endpoint is not required – rather, this pathway makes use of a biomarker or surrogate that is "reasonably likely to predict" a clinical endpoint
    - Importantly, the level of evidence required for regular and accelerated approval is the same
      - Efficacy and safety evidence will be based on two adequate, controlled, randomized trials that generate persuasive evidence
      - Exposure requirements also have to be met; this is especially relevant as almost all the drugs being considered for liver disease are chronic-use drugs
        - (Minimum) Exposure requirements by International Conference on Harmonization (ICH) guidelines: 1,500 people exposed to drug at doses that will be used clinically; 300-600 for six months and 100 for at least one year
    - Using small trials (e.g. a phase 2 trial) to obtain accelerated approval is not possible under FDA requirements, while it may qualify for conditional approval under EMA requirements
      - FDA and EMA/CHMP have more flexibility when it comes to rare diseases, for which smaller studies, sometimes with as few as 10 patients, may be sufficient

- Both accelerated approval (FDA) and conditional approval (EMA) require that the drug is being developed for an unmet medical need
- Accelerated approval makes a lot of sense for fatty liver disease and liver fibrosis because of the length of time it takes for disease to progress to a measurable, traditional clinical endpoint
- In general, it is up to the sponsor to justify the use of a particular biomarker for phase 2 trials
    - Validity considerations include whether it appears in the metabolic-physiologic pathway and whether it makes sense biologically
    - A good understanding of a drug's mechanism of action is very helpful
- The needed length of a phase 2 trial will depend on the specific biomarker: How long will it take to show an effect? OR for the biomarker to show a difference in dose-ranging or proof-of-concept studies?
- FDA is moving away from use of phase 1, 2 and 3 trials, and substituting "proof-of-concept" trial, "dose-ranging" trial and "trials designed to support marketing approval"

Following Lara's presentation, questions were posed to both Lara and Elmer and the other regulatory agency representatives present. The following section summarizes the ensuing dialogue.

- Question: How do data on target engagement or early proof of mechanism configure into the final approval process?
    - A well understood mechanism for drug action makes the approval process much easier
        - Jumping straight into clinical trials without understanding and/or proof of mechanism of action can be very risky
    - Drug development time can be shortened by rolling phase 2B (dose-ranging trial) patients into Phase 3 trial
        - Don't stop treating patients
        - Keep treating patients until the final report is submitted; correcting the dose if necessary
        - Continue the same patients into phase 4
    - Downsides to this approach:
        - requires appropriate statistical approaches (to account for alpha spending)
        - risks investing a lot of money into a plan that will not proceed, if phase 2 trial does not succeed
- Question: As we enroll patients earlier into the drug development process, are any concessions made with regard to length of toxicology studies (and when patients can be started on dosing/efficacy trials)?
    - There is no concession, but there is flexibility on the timing
    - Sometimes it is possible to submit longer-term toxicology data as they become available while already planning the phase 3 studies
- Question: Is it acceptable throughout the FDA that additional toxicology data for chronic dosing could be submitted while a protocol is ongoing or is that specific to this Division?
    - This is fairly acceptable throughout the FDA; although Divisions may differ on whether it is 30 or 90 days before dosing
    - Most Divisions realize you do not necessarily have to stop the patients while submitting the additional data
- Question: For first-in-class molecules, is a two-year study required to demonstrate negative toxicology and carcinogenicity?
    - Carcinogenicity studies only need to be submitted with the marketing application for most drugs, unless the drugs are high risk

- Question: To what extent and under what circumstance would the results of a Phase 2B trial, in a large enough number of patients, be considered sufficient for a registration package? What are the limitations? (Addressed to both Agencies)
  - This is addressed in the FDA's Guidance for Industry – Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products at: http://www.fda.gov/downloads/Drugs/.../Guidances/ucm078749.pdf
  - For the FDA, registration approval considerations would have had to be designed into the phase 2B trial design
    - Statistical plan for phase 2B would need to be adequate to use it for the registration application (so, it would need to be designed like a phase 3 trial with regard to statistical analysis plan)
  - For the EMA, there is a "one pivotal study" guideline which sets the conditions that would allow only one trial (pivotal or registration)
    - Depends on the strength of phase 2B evidence
- Question: Given the differences between the FDA and EMA processes and rules, what is the likelihood that both agencies will eventually speak the same language?
  - Because there are different governments writing the laws, that is not likely to happen any time soon
    - In the future, perhaps we could work toward a unified regulatory approach
  - It would be optimal for both agencies to have both provisional and accelerated approval processes and to clarify when/under what circumstances these can be used
  - There are instances of official harmonization of bureaucratic processes across global agencies, through the ICH, which includes Japan and East Asia.
  - There are ongoing, regular informal consultations between the agencies. At the request of applicants, parallel advice can be sought on use of biomarkers as surrogate endpoints or on the design of phase 3 trials (trials designed to support a marketing application).
  - The HIV Forum/Liver Forum, with its informal structure, allows people from different nationalities/sectors to hear each other's point of view and encourages the exchange of information and perspectives
    - Although we don't have the power, authority or capacity to change government law on either side of the Atlantic, the value of this Forum is that everybody can hear what everybody else is thinking and that each side can factor in those perspectives
    - Existing regulations are confusing, especially to start-up drug companies; this forum gives us the chance to educate and discuss how best to navigate the existing regulatory systems
  - With HIV, HCV and cytomegalovirus (CMV), we have found that we have been able to make progress without needing to re-write laws
- Question: When should liver impairment studies be conducted for pharmacokinetics and pharmacodynamics (addressed to both regulatory agencies)?
  - If the target population for phase 1 trials is healthy people, then liver impairment studies are not needed; but if the target population is patients with cirrhosis and the program is in Phase 2B studies, then liver impairment studies should have been completed earlier
  - In NASH pre-cirrhotics, impairment studies are probably not needed until the marketing application is submitted
  - The same would be true for the EMA
- Question: The current hepatic impairment guidance ranks patients who are Child-Pugh A, B or C as having mild, moderate or severe hepatic impairment (respectively), however we know that Child-Pugh C encompasses a huge range of clinical presentations. Do you foresee that companies will need to further elaborate on the status of Child-Pugh C patients with a score ≥10 when applying for approval for cirrhosis treatments?

        - Yes. The current guidance is inadequate. The FDA is currently asking applicants to stratify Child-Pugh C by MELD scores. We are requiring MELD scores for all patients with decompensated cirrhosis.
        - The EMA perspective is also (that the MELD score) should be available for the more hepatically impaired patients.
    - Question: What is your broad advice on companion diagnostics to evaluate treatment response? Specifically, what is the evidence needed to validate the companion diagnostic in the context of a clinical trial?
        - Discussion on companion diagnostics should be initiated early in the process and will involve input from FDA's Center for Devices and Radiological Health (CDRH) and the biomarker groups for appropriate advice.
            - The level of evidence is not as high as for a drug approval, but still can be quite high for the establishment of sensitivity and specificity of the diagnostic tool.
    - Question: Are there separate groups at the FDA and EMA that evaluate companion diagnostics or can these discussions be integrated into the SAWG or FDA meetings?
        - The FDA works together across divisions to evaluate diagnostics;
            - The Center for Drug Evaluation and Research (CDER) and the Division of Gastroenterology and Inborn Errors Products (DGIEP) help to coordinate that, but it should be brought up early in development
            - This issue becomes even more critical for device manufacturers who are developing NASH and fibrosis diagnostics, because it involves an entirely different part of the FDA
            - And, conversely, when devices and biomarkers come up for consideration, CDER needs to get involved early on. If CDER is not involved, sponsors should request it
        - In Europe, companion diagnostics are usually categorized as medical devices. They are certified (CE mark).
            - The utility of the diagnostic may be part of the evaluation process for a medicinal product, when it comes to marketing authorization, but it will not be a part of an evaluation of the device itself
    - An additional word on Special Protocol Assessments (SPA). These are declarations from the FDA that the design, conduct and analysis protocol for a clinical trial has been accepted for approval – essentially a legally binding contract where the FDA assumes the risk. These can be very useful for non-clinical studies (animal carcinogenic studies and product stability protocols), but can take a lot of time when applied to phase 3 pivotal trials. However, SPAs can be useful in achieving agreement between the sponsor and the Agency regarding the design and interpretation of study data submitted in support of marketing applications. They have the advantage of more or less guaranteed approval, if the trial was conducted successfully and no additional safety issues (not anticipated with the SPA evaluation) were encountered during development.

Bridge to next three sessions:

The next three sessions will focus on the three Working Groups: data standardization, definition of disease stages, and pediatric fatty liver disease/NASH. These three topics were identified during the first Liver Forum meeting as representing areas where there were significant gaps in the science. The Working Groups themes are interrelated. The strategy is to approach these questions from different angles and then integrate the discussions, as they mature. We will start with two brief presentations selected to inform our field, as we proceed with the Working Group processes: an example of a database collaboration leading to biomarker acceptance for PBC; and an example of industry/academic/FDA collaboration for more efficient assessment of ECG's for cardiotoxicity in clinical trials.

We begin with a note on biomarkers (slides 26-29). From Chris Leptak's presentation (Liver Forum 1), we learned that there are two basic functions for biomarkers. First there are those that mark the natural history of disease in the absence of therapeutic intervention. These can be further differentiated into diagnostic and prognostic biomarkers. Then there are those that measure the response to therapeutic intervention. These include predictive biomarkers, measured prior to a therapeutic intervention, pharmacodynamic (PD) biomarkers, which measure a biologic response to a therapeutic intervention, and efficacy response/surrogate biomarkers, which substitute for a clinical outcome.

- We need to be clear on the specific function of each of the biomarkers that we take under consideration
- We also need more standardization (of measures and definitions) before we can address biomarkers that measure the response to an intervention (see slide 33 in http://www.hivforum.org/storage/documents/2015/Liver_Forum/miller_presentation_liverforum2.pdf)
- Most importantly, unless we can clearly define the baseline, it will be difficult to measure outcomes

Session #3: Toward Data-driven Collaboration
Moderators: Veronica Miller and David Shapiro
Discussants: Bettina Hansen, Borje Darpö, Joanne Imperial, Rohit Loomba, Rob Myers

- Bettina Hansen presented an overview of The Global Primary Biliary Cirrhosis (PBC) Study Group. She discussed lessons learned from studies of surrogate markers in PBC trials (Slides available here: http://www.hivforum.org/storage/documents/2015/Liver_Forum/hansen_presentation_liverforum2.pdf)
    - Impetus for the PBC project: there is only one drug available for treatment of PBC and some patients respond sub-optimally, so new treatment options are needed
        - New treatments are being developed, requiring consensus on surrogate markers
    - Like NASH, PBC is a slowly progressing disease, but in contrast to NASH, it is a rare disease
        - Most patients present with early disease symptoms
        - A trial would take too long (8-10 years) to complete if dependent on hard clinical endpoints, so a surrogate endpoint was needed
            - For shorter study duration
            - To bring drug to market quicker, which is especially important to the patients not responding to currently available treatment
        - A number of different surrogates were proposed, encompassing different response criteria (see slide #5 in presentation)
    - A good surrogate endpoint should
        - measure change (in value) reflective of changes that would be observed in the true clinical outcome/endpoint
            - capture the whole effect of the intervention
        - be noninvasive and easy to measure
        - precede the clinical endpoint and is assessed within a short time frame
        - easily and quickly detect danger of toxicity
        - lead to shorter study duration
        - encourage study participation and therefore influence potential sample size
        - reduce the time and thereby the cost of the study
        - have already been proven to accurately reflect a clinical endpoint
            - recommend using meta-analysis of both surrogate and clinical endpoints as in clinical trials of multiple related drugs, if available
    - Reasons that a surrogate endpoint often fails during validation (see slide #8):
        - Not in the causal pathway of the disease
        - Of several causal pathways, the intervention only affects the pathway measured by the surrogate
        - The surrogate is not in the pathway of the intervention's effect
        - The intervention leads to its own adverse effects on the clinical endpoint
    - In order to establish the validity of a surrogate:

        - Need in-depth understanding of disease process and mechanism of action of the intervention
        - Need FDA/EMA agreement on the use of the surrogate
    - A four-level hierarchy for validating and approving surrogates has been established (see Fleming TR, *Health Aff*, 2005 and Fleming TR, Powers JH, *Stat Med*, 2012):
        - Levels 1 and 2 are likely primary endpoints in registration trials
            - Level 1: a true clinical efficacy measure
            - Level 2: a validated surrogate
        - Levels 3 and 4 might be considered as primary endpoints for accelerated approval in clinical trials:
            - Level 3: a non-validated surrogate considered reasonably likely to predict clinical benefit
            - Level 4: a correlate that is a measure of biological activity, but not yet shown to predict clinical outcome
    - Background: The pharmaceutical industry provided the impetus for the Global PBC Project, when they suggested that industry and academia join forces to prove surrogacy
        - The aim of the project is to determine the prognostic significance of ALP and bilirubin as surrogates for transplant free survival
        - First meeting took place in Berlin, 2011
        - Investigators who had already produced response criteria were invited to contribute their data to one combined database for purposes of meta-analysis of patient-level data
        - Study protocol, consortium agreement, case report forms and letter of expected inclusion were developed
        - IRB approval was obtained separately for each participating institution
        - Contributing centers include those with both retrospective and prospective study data
            - Face-to-face meetings are held twice a year at AASLD and EASL
        - Site visits are regularly scheduled for PBC Project staff to provide technical assistance with data collection and quality monitoring.
        - 15 centers have joined the consortium and approximately 6,000 patients are in the database
            - With a sample of this size, it is possible to do adequately powered subgroup analysis (e.g., males with PBC)
    - Database consists of the two hard endpoints (death or liver transplantation) and surrogate endpoints (i.e. biomarkers) measured at baseline, one year, two years and at each follow-up visit through 2012
        - A very large database has been amassed representing 40,000-50,000 total visits by 6,000 patients
        - In addition to ALP and bilirubin, other clinical lab data have been collected, including: ALT, AST, and decompensation and hard endpoints, if they occurred. (see Case Record Form, slide #16)
        - 85% of Global PBC Project patients had been treated with UDCA
    - Governance consists of a steering committee, whose functions include:
        - Review of proposals submitted by participating investigators
        - Setting up writing committees, approval of author lists (in advance)
        - Criteria for co-authors: involvement in design of study; contribute sufficient number of patients, lab data, materials; perform interpretation of the data; participation in the drafting of the article
        - Two types of papers generated: general papers involving the entire dataset and local papers on small issues
        - Collective ownership of data
    - Measures of success:
        - Continuous growth and new centers requesting to join the consortium
        - Six publications, with two more in process

        - PR and networking are priorities: regular investigator meetings, newsletters, website and new risk score calculator
        - Scientific presentations made yearly at EASL and AASLD
        - Grant applications coming in; PhD student applicants
    - Challenges:
        - Obtaining financial support for PhD students—mostly from the pharmaceutical industry
    - Summarized answers to participant questions follow:
        - This is entirely an investigator driven initiative (despite some grant money having been obtained from the pharmaceutical industry)
        - The meta-analysis was not of the conventional sort; it did not amalgamate data from multiple published studies, but rather amassed previously collected data from multiple clinical sites and observational studies and then "prospectively" harmonized the data, so that project could go forward with a standardized database
        - With regard to potential, future Liver Forum projects, this type of endeavor could be made easier by designing studies collaboratively from the beginning
        - Patients also could be enrolled prospectively
        - It should be noted that biochemical changes are the principal biomarkers of PBC progression and biopsy is no longer routine for PBC, including for diagnosis. In contrast, biochemistry is not the principal method used to track disease progression in NASH; clinical outcomes are. The biochemistry is unlikely to be as helpful. Biopsies are still the standard, which presents its own conundrum
        - NASH is similar to HCV insofar as progression to cirrhosis is the major clinical outcome assessed
        - It also should be noted that since all the case data came from tertiary care academic medical centers with transplantation units that there could be selection bias towards the more severe cases
        - Patient data were not collected from pharmaceutical companies for the PBC Project.
        - The purpose of this Liver Forum Working Groups is to start planning now on collaboratively designing trials and prospective studies in such a way that data are standardized and shared and where pooled analyses (and meta-analyses) can be performed

- Borje Darpö made a short presentation on standardization of ECG data from the Cardiac Safety Research Consortium (CSRC), an organization which facilitates collaboration among industry, academia and the FDA (with representation from EMA and PMDA). Borje Darpö co-chaired the Scientific Oversight Committee of the CSRC for the last three years. (slide presentation may be found here: http://www.hivforum.org/storage/documents/2015/Liver_Forum/darpo_presentation_liverforum2.pdf)
    - ECG abnormalities, in particular prolongation of QTc interval are common among patients with advanced liver disease (specifically among cirrhotic patients).
    - Commonly used correction methods have led to inappropriate exclusion of patients from studies
        - Results in QT prolongation values which are too long for patients with elevated heart rate, making it appear that trial drug causes this effect
    - As such, standardizing the use and interpretation of machine readings of ECG data would be advantageous to NASH researchers
    - Two approaches are taken to reduce the exclusion of patients from studies based on ECG abnormalities:
        - Defining the study drug's effects on ECG earlier in development using Exposure-Response (ER) modeling
        - Efficient and accurate "alert-triggered" evaluation of screening ECGs
    - The first approach was used in a IQ-CSRC study looking at QT interval changes across a range of increases in drug plasma levels in a phase 1 setting
        - They looked at 5 positive drugs and one negative drug (with regard to their effects on ECG) with prespecified criteria for determining whether an effect was present or not

  - This was a small scale study, but very successful as an exercise in collaboration across industry, academia and regulatory sectors
  - A second approach would be to use an alert-triggered method whereby screening ECGs that demonstrate abnormalities (as identified by machine or by the investigator) triggers an evaluation at the central lab, making sure that patients are not inappropriately excluded
    - The alert-triggered approach is cost-saving compared to a central core lab looking at all ECGs
    - In a collaborative study, a significant number of ECGs triggering an alert were found to be in the normal range following core central lab analysis yet these patients would have been excluded from the trial based on the read-out alone

- Data Standardization Working Group Overview [Joanne Imperial and Andrew Muir, co-chairs]
  - The purpose of this Working Group is to discuss the types of baseline data we need to collect and whether these data can be standardized with the intent of pooling data at some point for large scale analyses
  - Critical Issues:
    - The population affected by NASH is heterogeneous and natural history is not well described
    - Spontaneous reversal of disease may occur in a substantial number of patients
    - Liver biopsy still widely used as diagnostic and prognostic biomarker
    - Outcomes may be very different for patients with any given (baseline) level of fibrosis
  - Objectives:
    - Standardize the definition of the target population before entering into trial
    - Identify biomarkers that will assist in better defining the population and use those in selecting target patients, so that some of the heterogeneity can be eliminated
  - Standardizing patient populations: looking at ways to pool data and at possible future trial designs
    - How should patient groups be stratified: Sicker versus less sick? Pre-fibrotic versus fibrotic?
    - Different endpoints will need to be defined based on the target population, their baseline measures and the outcomes studied
    - It may be advantageous to look at patients with more advanced disease because the endpoints (death, transplantation) are more well defined, the likelihood of them getting better is likely lower and it will be easier to study their natural history
      - Using liver transplantation as an endpoint is problematic because there are many reasons, other than the primary underlying disease, that a patient would not get a transplant
      - MELD is a better way to evaluate the need for transplant
    - Control group patients can be used to elucidate natural history of NASH
    - Many biomarkers and assays being considered
    - How to identify more rapid progressors from those who progress more slowly?
  - How to move forward with collaboration: What mechanisms need to be put in place so that all the players are able to participate in a comfortable way?
    - Confidentiality agreements
    - Identifying what types of data can be shared
    - Rules for use of data; pooling data and establishing a large database (from which we can learn more about natural history and which can be used to guide us in the identification of appropriate endpoints to use in accelerated clinical trials)
  - Another potential area for collaboration would be to determine the possibility for combining agents from the different pharma companies and developing clinical trials to test the impact of drug combinations

Data Standardization Working Group: Panel Discussion:

- Regarding the fastest and most efficacious way to standardize and pool data and establish a natural history cohort
  - A specific set of required baseline characteristics needs to be established
  - We can try to use/adapt the model developed for PBC for NASH
  - There are three broad categories of patients to be treated: [steatohepatitis/no fibrosis, early fibrosis, advanced fibrosis/cirrhosis];
    - future studies/trials should be designed to address these target groups as separate "pods"
  - Within each pod, patients can be further stratified by specific, appropriate risk factors (e.g. MELD and QTc for cirrhotic patients) or subgroup cirrhotics into compensated versus decompensated and pre-cirrhotics into fatty liver disease, NASH and NASH with fibrosis
  - Spontaneous regressors need to be identified otherwise they will prove detrimental to discerning a treatment effect (there may be as high as a 40% spontaneous regression rate, based on previous studies)
- Resolution of disease also needs to be defined, including what measurements (surrogate endpoints, biomarkers) can be used
- Role of imaging needs to be clarified; how can other functional tests be utilized
- Data could be pooled now, incorporating existing/completed studies by investigators present; there may be some restrictions on pooling data for certain trials, for example those sponsored by NIDDK, which first need to go into public domain
- Once data are acquired they can be anonymized; these data could be used to establish sample sizes for specific outcome studies. The Liver Forum has a role to play here
- Static versus dynamic markers: ideally we would have a baseline test that would predict patient outcomes; however we should consider looking at multiple tests over time, as did the PBC Project
- To detect clinically significant events, short duration treatment studies may need to be followed with open label extension periods
- (FDA perspective): open-label extensions won't help in the context of accelerated approval, it will only generate safety data
  - For efficacy data, we need to think of a controlled extension trial (e.g., extend placebo control) or, hopefully in the future, have the appropriate historical control cohorts in place
- Ethical issues must also be addressed with regard to longer term retention for accelerated approval trials, if it is shown that outcomes are different between treatment and control group
  - The Liver Forum can make a major contribution by establishing a database of historical controls for this purpose: then you do not need to retain placebo control after marketing applications are submitted
  - If you outperform a historical control by a significant margin, that demonstrates a drug effect

Session #4: Speaking a Common Language: Disease Definitions
Moderators: Markus Peck and Arun Sanyal
Discussants: Sophie Megnien, Stephen Harrison, Laurent Castera, Jude Oben, Rebecca Taub

Overview: The global strategy for this Working Group is to come up with consensual definitions related to the disease as a whole. Specifically, it will summarize the current state of the science and identify gaps in knowledge, opportunities for advancement/filling those gaps and opportunities for collaboration. In stepwise fashion, the group will

- work to define NASH
- establish criteria for the diagnosis of NASH
- identify flaws or weaknesses in the methodologies currently in use to diagnose and track disease progression

- develop strategies for identifying patients at risk for disease progression and
- identify clinical trial endpoints

The Disease Definitions Working Group and the Standardization of Data Working Group have significant overlap, but this is intentional. By approaching the problem from two different angles, it is expected that the output of these groups will converge in the end.

During the first conference call held by the Disease Definitions Working Group, the state of the science was reviewed, including identifying the tools available and the gaps or flaws of these tools. The task now is to identify the terms that need to be defined, clarify ambiguous terms and those with overlapping definitions, and validate a quality set of terms that can be used to diagnose and track disease, using existing/available data. Considerations include: is the definition valid, reproducible, analyzable from a statistical point of view?

Panelists and participants took note that pathologists and radiologists are missing from the Liver Forum at present and should be included in this process.

The following is a synopsis of a broad ranging discussion that focused on the definition of NASH and on the methods/criteria used to diagnose and track the progression of NASH (and NASH related fibrosis) in treated and untreated patients. The overall purpose of the discussion was to spur the development of criteria that can be operationalized across multiple trials and which will allow the regulators to quantify results and readily compare one study to another.

Refining and standardizing the definition of NASH was the lead topic discussed. Will it come from histopathology or from biochemical markers, or will it be radiographic? Ultimately, it probably will come from some combination of non-invasive tests. Right now, the definition of NASH rests on liver biopsy, but is not based on histopathology alone. NAS score is also used. Current practice guidelines from EASL and AASLD rely on the hallmark feature of ballooning in the setting of steatohepatitis (including lobular inflammation, more than 5% fat and in more advanced disease, portal inflammation) for defining NASH. The historical paradigm held that NAFLD does not progress, while steatohepatitis progresses; therefore we need to better define steatohepatitis.

Nomenclature was another issue raised. The terms currently used are confusing to the patient and may also lead to misdiagnosis. The name alone, non-alcoholic fatty liver disease and non-alcoholic steatohepatitis are defined by the absence of a clinical risk factor. Furthermore, it is well known that many cases of so-called NASH are in fact influenced by a significant history of alcohol consumption. Patient self-report of alcohol consumption may not be reliable and differentiation between alcoholic and non-alcoholic liver disease is problematic; sometimes it is a combination of both. Furthermore there are different presentations of NASH in regions throughout the world, where a substantial proportion of patients are not obese or overweight. NAFLD/NASH is a disease that affects populations across the lifespan, including pediatric patients; therefore a name that is appropriate for all age groups was suggested.

This raises the issue of etiology. A model for the differential diagnosis of NASH was proposed. The multicomponent framework includes: 1) two phenotypes: steatosis or steatohepatitis; 2) stage, as established by liver biopsy: no fibrosis or fibrosis, which is conventionally staged; and 3) etiology: non-alcoholic, mixed non-alcoholic and alcoholic and/or possible underlying genetic condition that predisposes to fatty liver disease. The first step is to identify the underlying pathology as hepatic, steatohepatitis or cholestatic. If a steatohepatitis pattern is seen, then the next step is to arrive at the etiology, informed by clinical data. Primary NAFLD would be identified, when there are no secondary causes for hepatic steatosis. Related to the issue of etiology is the notion that NAFLD/NASH is the liver component of pre-existing, predefined metabolic syndrome. It is on this basis that EASL has recently issued a new guidance.

And this in turn leads to the issue of histology. NASH is a histological syndrome characterized by a liver phenotype. But many influences can contribute to the syndrome and there is still the possibility that NAFLD represents several different underlying diseases.

Liver biopsy is currently the method used to establish a diagnosis and will remain so until we have noninvasive tests that provide a reliable marker of disease. That said, results of liver biopsy alone do not suffice for determining disease, as a substantial proportion of patients with NASH defined by liver biopsy and a NAS score of 4 will not progress. Determination of fibrosis stage from liver biopsy remains the clinical endpoint used in drug development. It is also the (imperfect gold) standard that any new biomarker (or liver disease measurement tool) will need to be compared to. The obvious goal is to replace the biopsy with non-invasive biomarkers, as liver biopsy cannot be used for ongoing patient monitoring. Patients that have NASH with fibrosis are likely to be the target population for future drug trials because this subgroup of patients has measureable disease, the beneficial effects of drug can be quantitatively measured, and they are also the group most at risk for disease progression from NASH. An alternative point of view, also expressed here, holds that fibrosis should not be part of the disease definition because it represents a long-term outcome. Alternative methods are needed for determining which patients are at high(er) risk for progression, while in earlier stages of the disease.

From a regulatory perspective, NASH needs to be defined in a manner that has content and face validity. Standardization is key. The criteria used should be objective, quantifiable, analyzable and reproducible from person to person. Then we must define the population within NASH that is “worthy” of therapy (and thus the target population for clinical trials). NASH with fibrosis has the worst prognosis and this is the group the Agencies are interested in targeting therapy towards. The goal of therapy then would be to either eliminate NASH or improve fibrosis, without worsening steatohepatitis. Using that as a framework, an effective treatment would need to get rid of ballooning and either reverse fibrosis or demonstrate that it did not progress. This definition assumes that there is some fibrosis at baseline. At present, regulatory agencies would require liver biopsy at both the beginning and end of a trial.

Throughout these discussions, there has been some confusion between defining the disease and defining the target population(s). This partly is due to the fact that various clinical subgroups exhibit the same histology. Again, from a regulatory point of view, it is important to consider whether all or just some of the subgroups should be included in a trial. This would depend on the trial drug’s mechanism of action and with which subgroup the drug is expected to work. Furthermore, risk of disease progression will also factor into the determination of which patients are chosen for inclusion and the endpoints selected for use in a trial.

- It would be difficult/tricky to come up with totally new definitions without conducting clinical studies that support and demonstrate that those definitions are valid; rather than “reinventing the wheel” we should “harmonize” the existing definitions

- In order to arrive at these definitional objectives, a grid that encompasses all the parameters currently used in operational definitions of steatohepatitis was proposed. This would be culled from the published literature and then used as a basis for comparison and evaluation.

- Echoing the discussion from the Standardization of Data Working Group, a collaborative, multi-site natural history study was endorsed, one that is modeled on the data collection framework developed by the Global PBC Project

- A recommendation was made that the Liver Forum also conduct a literature review focused on the prognostic relevance of fibrosis stages in NASH and on defining (the role of) fibrosis in the context of necroinflammatory activity and bridging fibrosis

In summary, the Working Group was charged with developing an operational definition of the disease and identifying the subset of the patient population at highest risk of disease progression and therefore

the most likely target group for a clinical trial. Secondarily, it needs to establish criteria for validating new technologies for biomarker development.

Session #5: Pediatrics and Adolescents
Moderators: Carol Brosgart and Gary Burgess
Discussants: Joel Lavine, Krishna Polu, Miriam Vos, Dennis Grasela, Stefan Neubauer

This session opened with a presentation by Ruby Mehta, from the FDA: Typically, in the drug development process, children are enrolled later in the clinical trials, after safety and efficacy has been established in the adult population. If the mechanism of disease in the pediatric population is the same as it is in adults, then an extrapolation or partial extrapolation method can be used in establishing a pediatric indication. In NASH and NAFLD trials, the agency is still struggling to determine whether the pediatric disease is mechanistically similar to adult disease.

Joel Lavine and Miriam Vos followed with their observations on pediatric NAFLD/NASH. Joel Lavine began the discussion with a more detailed description of FDA experience with clinical trials in children. He reviewed the history of the FDA Modernization Act of 1997 and the Best Pharmaceuticals Act, which followed. These were intended to result in pediatric research equity and to allow a six-month patent extension for companies that took the extra steps necessary to establish a pediatric indication. Despite this history, most companies request waivers and most often waivers are granted. This puts the child and the family at risk. With regard to NAFLD/NASH and the high prevalence of disease in children, forward thinking is that new drugs will have to have pediatric labeling and that there will be pediatric trials.

The following issues were identified with regard to pediatric liver disease and clinical trials in children:

- Much less is known about the natural history of the disease in children than in adults
- There are significant and unique challenges in doing clinical trials in the pediatric population, especially with regard to the use of liver biopsy and other measures that could be viewed as disincentives to participate, including blood draws, imaging, pilling and dose adjustments (necessitated by rapid body weight changes that take place in children).
- Ethical issues arise with regard to obtaining informed consent
- There are benefits, too:
    - Children are a captive audience.
    - They can be said to have NAFLD or NASH in its “native” state, without the influence of alcohol.
    - Parents facilitate participation and compliance
    - There is a good record in NIH-sponsored NASH Clinical Research Network of children completing randomized controlled trials with biopsies at the beginning and end of trial

We do not know whether the etiology of disease is different in children than it is in adults, nor if there is a change in disease manifestations as children approach puberty. That children are so young when they are found to have NAFLD/NASH implies a genetic susceptibility, an environmental susceptibility or both. Various phenotypes are seen in children, just as in adults. NASH is seen in very obese children and in thin children with a family history. Cases of compensated cirrhosis have been seen in children as young as eight. Children generally do not go into end stage liver failure or need transplant before they reach age 17, at which point they are lost from pediatric studies, by definition, and there is no national database to track these individuals into adulthood.

Demographics also play a role in defining this epidemic in children. Boys outnumber girls and indigenous Americans seem to be particularly susceptible. Children in other parts of the world likely have unique susceptibilities as they are developing disease despite low to normal BMI.

Because of the reluctance to biopsy young children, it can be assumed that the pediatric population is underdiagnosed. Blood sampling, imaging and other monitoring procedures also dis-incentivize participation. These factors, combined with the reluctance of parents to enter children into a trial, make it

hard to power studies adequately. While it is an added burden to involve a family in study visits, parents can also aid a study by ensuring compliance. There are also issues related to risk and whether placebo controlled trials are appropriate in children. Pediatric trials will not be testing first-in-class drugs because of their unique risk. And pharma has been less interested in the pediatric population because of the added concern about adverse outcomes. For these reasons, it is important to be realistic in making sample size goals for pediatric trials.

Several issues were raised with regard to the use of biopsy in children. Inclusion criteria may specify pre-existing NAFLD or NASH, but that diagnosis is hard to establish without a liver biopsy. Consensus on what constitutes standard of care biopsies for children has not been established; guidelines from ACG, AGA, AASLD to not provide sufficient direction. Thus it would be hard to identify potential cohorts from existing clinic populations. Furthermore, liver tissue from a child with NASH looks different from that in an adult. A pediatric biopsy specimen may have portal fibrosis and portal inflammation, but it will generally lack lobular inflammation or perisinusoidal fibrosis. And it will lack ballooning. These signs do not indicate benign disease, as these patients may develop bridging fibrosis, have a high ALT and/or have diabetes.

The question was raised as to whether children should be stratified by age group or by the onset of puberty, because the biology of the disease following puberty is different. The point was made that children younger than eight are unlikely to be enrolled in trials, in part because of the volume of blood required. If older children are stratified by puberty status, should they be further subgrouped into those who have clear-cut disease and those who have mild or borderline disease? A natural history study, targeting these age groups, would help clarify these differences.

From a regulatory point of view, stratification by age would be helpful, given the physiological changes that occur at puberty. With (current) NASH trials, the FDA has seen surges in insulin resistance and growth hormone with puberty. An additional issue: the FDA evaluates all pediatric protocols with regard to its potential to directly benefit the child. The FDA and EMA are also concerned that children participating in trials face no more than minimal risk. Also, under the Pediatric Research Equity Act (PREA), drug companies must develop age-appropriate formulations. Because of PREA requirements, it is advisable to submit a pediatric study plan by the end of the Phase 2 trial (if not earlier, to satisfy EMA standards). The regulatory agencies would also want to see the differences between adults and children properly characterized and to have biomarkers validated in the adult population, before proceeding with their use in children. Finally, if a clinical benefit trial establishes that a surrogate marker works as an endpoint in adults, and you are planning to use that surrogate in a pediatric trial, it may be justifiable, from the regulatory standpoint not to conduct a clinical benefit trial. It is critically important to add imaging and other biomarkers to the compendium of measures, not just for establishing the results of a clinical trial, but for ongoing monitoring of patients.

Finally, because it is entirely possible that NAFLD/NASH represents a continuum of disease that often begins in childhood, and that segregating those with pediatric interests/expertise into a separate Working Group artificially divides this group from the other Working Groups, some members of the pediatric Working Group should join the definitions and biomarkers Working Groups, so that there is overlap of effort.

Liver Multiscan

Stefan Neubauer gave a presentation on Liver Multiscan, a new MRI-based imaging technique. (The slides and technical details for this presentation can be found here: http://www.hivforum.org/storage/documents/2015/Liver_Forum/neubauer_presentation_liverforum2.pdf)

Multiscan is a multi-parametric tool to characterize liver disease developed at the University of Oxford Clinical Magnetic Resonance Research Center and by Perspectum Diagnostics. It measures three things: 1) iron content, 2) fat fraction and 3) inflammation and fibrosis, which is novel, and quantifies

these in a liver inflammation fibrosis score (LIF), using a new and patented MRI-based method (described in detail in the slides). No additional hardware is required. The AUROC is 0.94 (CI 0.89-0.99) to detect any liver disease in a general population. The whole liver can be assessed or just one slice. It works well in obese patients and in those with ascites. The procedure is brief, taking about five minutes. A color coding system is used to indicate Ishak score. There is a high correlation with fibrosis score obtained from histology slides and LIF. The method detects ballooning well. It has high rates of inter-rater reliability and power calculations indicate that relatively small sample sizes would be needed to detect an effect size corresponding to a LIF score change of 3 to 2, indicating a change from moderate to mild disease, in a clinical trial setting. These results were obtained in clinical testing in a highly specialized MRI research center. They are now undergoing confirmation in "real world" settings in Birmingham and Edinburgh.

Opportunities for Collaboration on Pediatric Trials. Several specific opportunities were identified: 1) If in the context of conducting meta-analysis of existing studies, endpoints most helpful in children are identified, companies could be encouraged to use the same endpoints and similar trial designs. 2) Sharing techniques for micronization of assays would be very useful insofar as the blood volume required could be reduced. Blood samples will be needed to understand pharmacokinetics as well as for clinical biomarkers. 3) Checkerboarding of some of the serum biomarkers across clinical trials would make it possible to avoid sampling each child at each clinic visit for each of the biomarkers. This implies using a population pharmacokinetics approach, looking at the data together and looking at them longitudinally. 4) It would be very useful to build a disease model for NASH that identifies the different phenotypes of pediatric disease. Understanding the etiology of the disease is paramount for the drug companies to match their drugs with the appropriate patient group. 5) Industry collaboration on sharing of data for placebo controls would give the field at large a better understanding of event rates, spontaneous regression and some of the other hallmark characteristics of the disease.

There are differences in regulatory approaches to pediatric drug trials between the FDA and the EMA. The EMA has a separate pediatric committee (PDCO), whereas this is not the case with the FDA, where pediatric expertise is brought into the general medicines approval process. PDCO is responsible for reviewing and approving plans for pediatric clinical trials, however the CHMP is responsible for final approval of drugs, including those used in the pediatric population. The Liver Forum is in the process of requesting representation from PDCO on the Pediatric Working Group.

Session #6: Summing Up and Next Steps
Moderators: Gary Burgess, Veronica Miller and Arun Sanyal

The moderators expressed their thanks to AASLD and EASL for their ongoing support of the Liver Forum as it moves forward as an independent entity focusing on regulatory development. They also expressed thanks to the FDA and EMA for their participation in and support of Liver Forum activities. The combined support of these organizations gives us strong anchoring on both sides of the Atlantic and the Forum will continue to reach out to colleagues on both sides for their assistance with specific projects. The larger meetings will continue to alternate between Europe and the U.S. We will conduct a survey to obtain your feedback about this meeting and suggestions for future meetings and Working Group projects.

Our goal is to maintain the balance of participation from people on both sides of the Atlantic and to be as inclusive as possible.

Following this meeting, the Working Groups will become increasingly active. Conference calls for each group will take place approximately every six weeks. Minutes from those discussions will be circulated. Timelines and deliverables will be established for each group. At the next Forum meeting, we expect that presentations will reflect significant progress made toward accomplishing each product. Time is crucial as we need to place these products in the public domain for everybody's benefit.

We are also considering whether to convene a separate diagnostic roundtable, over the course of the coming year, to enable some of the developments in diagnostic science to be discussed more fully. Liver Forum 3 will be held in the fall of 2015, preceding AASLD, in San Francisco.