OSIC Meeting Documents | ATS 2018 | May 18-23rd

[OSIC Meeting with Brigham and Women's](#BW)

[OSIC Meeting with NJ Health](#NJ)

[OSIC meeting with Pulmonary Fibrosis Foundation](#PFF)

[OSIC meeting with Royal Brompton](#RB)

[OSIC meeting with University of Vienna](#UV)

[OSIC meeting with American Lung Association](#ALA)

[OSIC meeting with the University of Iowa](#UI)

**OSIC Meeting with Brigham and Women's**

**Attendees:** Roberto, Raul, Megan

**Action Items:**

* OSIC to create legal entity and structure
* Raul to send representatives for Data Working group etc. to Roberto and Megan
* OSIC to set up infrastructure for WGs
* Raul will talk to Data approval committee to educate them around OSIC and find out timeline (once structure and standard is in place)
* Roberto to send presentation to Raul

**Notes:**

* Individual conversations with all data contributors
* OSIC is close to getting the legal entity created
* OSIC is starting several working groups around the legal framework, and model that the institutions will need to donate the data
* Data working group to define:
  + The purpose of the data / access controls
  + How the data will be compiled
  + How the data will be stored
* Four-year project plan included 3 kaggle challenges. Need to continue to add to the database
* Founding members will have representatives on the Board and in working groups
* Hiring an Executive Director
* Once Data Group agrees on standard, the Privacy group needs to work on risk of re-identification etc.
* Each institution should nominate someone to be their voice in each working group
* Data will be compiled, encrypted, sent to vault, anonymized by a third party, then used by OSIC
* Two options for access
  + Data contributing members only access
  + All members access
* Concern with this is that anyone who has Kaggle challenge will have access to that data, but they will be required to sign statement
* Process and timelines:
  + Nominate some people to participate
  + After legal entity, contract will be signed by OSIC and member
  + Working groups create model and standard
  + Begin to compile data
* Data approval committee (high leadership sits there)
* Address the black-market issue
* OSIC could become a model for other medical industries
* ILB’s in place
* Legal team will be a separate step

**OSIC Meeting with NJ Health**

**Attendees:** Roberto, Stephen, Elizabeth, Megan

**Action Items:**

* Stephen will volunteer for Data WG
* NJ Health to let OSIC know the expense (flat fee) and the steps involved in collecting the data
* OSIC creating infrastructure for working groups in two weeks
* OSIC creating legal entity in one month
* OSIC creating Data WG
* OSIC to figure out if we can pay institutions through PFF, and collect data through PFF

**Notes:**

* Roberto walked through the model, including the Data WG, data compilation, curation, protection and ethics
* WG1 – Data, WG2 – Ethics
* NJ scans may flow through PFF
* NJ will do internal anonymization
* Could store on hard drive or send electronically
* The data group needs to define the standard data fields that would be needed. Once all processes and policies are defined, then can begin the data collection process.
* 2 months to set the standards
* Concern: effort of compilation. OSIC can provide support to help compile the data.
* We need to know how much money and the mechanism to collect
  + Per exam cost – NJ to relate flat fee to OSIC
  + Query the Pacs (to run after work day), use medical record #, need to be cleaned up to just include the right data, clear redundancies, visual QA,
* OSIC can’t pay per scan, but as donation to the institute for coordinator support
* May be easier to go through PFF to pay
* OSIC will hire 3rd party to anonymize and standardize the data
* 700 scans

**OSIC Meeting with Pulmonary Fibrosis Foundation**

**Attendees:** Roberto, Elizabeth, Rex, Greg, Bill, Megan

**Action Items:**

* PFF to find a rep for the Data WG and send to Roberto and Megan – will it be Greg Cosgrove?
* OSIC creating infrastructure for working groups
* OSIC creating legal entity
* OSIC creating Data WG

**Notes:**

* Looking for 1500 scans by the end of 2018
* Built infrastructure, but need help with nuances, particularly with data transfer.
* OSIC can work with individual orgs, but would prefer an aggregator, if PFF can fill that role.
* Data WG will be working on data protection, curation and compilation
* Stages include data compilation, encryption, transfer to vault, 3rd party anonymization, transferred to Data Management platform for use by OSIC
* Concern not that there will be a leak, but if there could be reverse anonymization.
* Registry is enrolling. Sabre was doing it but failed technically. Upload anonymization was not satisfactory.
* Last few days trying to move process forward. Need agreement between Sabre and imaging group
* Should be 6000 scans from 2000 patients
* Will have 4000 – 6000 scans to contribute but will need coordinator help to compile the data
* Need to encourage those sites who enrolled to agree.
* The scans are in 40 different databases.
* 6 months seems reasonable to acquire the 1500 scans
* 1961 in the registry
* Encourage the coordinators to continue instead of enrolling patients to focus on inputting the data
* Goal is to acquire similar data
* Siemens can offer anonymization tools
* Looking to have a minimum of 3 challenges
* We are mapping people who are passionate about different topics
* Need to define the standards in the Data WG
* Early and big win for the foundation.
* Big splash – PR on the wire etc.
* PFF – only issue, cost recovery. OSIC try to mitigate it.
* Rex can help with technical aspect, along with Kevin Flaherty and Greg Cosgrove

**OSIC meeting with Royal Brompton**

**Attendees:** Roberto, HB, Justin, Zheng, Norbert, Megan

**Action Items:**

* Toby to find a rep for the Data WG and send to Roberto and Megan
* OSIC creating infrastructure for working groups
* OSIC creating legal entry

**Notes:**

* One concern is traceability of publications
* By contributing don’t want to prevent their own studies, or other people to use the data.
* Will discuss this in Data WG
* Membership agreement will follow after legal entity is established – one month
* Data WG will include data curation, data protection, data compilation
* Data WG will create data set requirements and access controls
  + Option 1: only data contributors will have access to the database
  + Option 2: all OSIC members will have access
* Orgs will donate the scans and rights to use them
* Providing scans as a resource for academic development of open source tools is fine. Commercial organizations having access without feeding value back to OSIC is more challenging.
* Early access to database will be the member benefit
* If a member develops an algorithm, are they required to be open source? As of right now, no.
* Royal Brompton potentially has 10,000 scans
* Important to reduce barriers to get funding
* In Kaggle challenge will have access to data
* If a company wants to participate, they would contribute the minimum amount of scans – if they have them
* Different types of members with different resources.
* EU data center, the rest to US
* Toby represents Royal Brompton and Imperial College
* Importance of race and region? Doesn’t influence decision making, so makes sense to delete that info if it makes sense. The fewer data points the better.

**OSIC meeting with University of Vienna**

**Action Items:**

* Need data requirements ASAP for legal approval – WG to decide on
* Helmut to begin legal process
* Look into how to initiate payment for the data compilation resources
* OSIC creating infrastructure for working groups
* OSIC creating legal entity

**Notes:**

Data Working Group

* Professor Langst will represent Vienna in the Data WG
* Data WG will agree on how anonymization will take place, and who will have access
  + Option 1 – only data contributors will have access to the database
  + Option 2: all OSIC members will have access to the database even if not bringing scans
* If we go with Option 2 – it will be easier to get money
* Next two weeks, nominate reps from each “member” for the data wg
* WG will define the data set

Technical process details

* Server is located in the EU
* 2 entities, Vienna general hospital, medical univ of Vienna
* Need to get approval from both legal departments
* Can begin data compilation now
* Database will be ready at the end of the summer, but if need data sooner, it can be done manually
* OSIC can provide a small amount of money for this effort
* Need to look into how the money can be transferred
* 300 scans

**OSIC meeting with American Lung Association**

**Attendees:** Elizabeth, Norbert, Zheng, Roberto, Al, Harold, Megan

**Action Items:**

* ALA to find a rep for the Data WG and send to Roberto and Megan
* Al to find the right contacts with radiology and pacs technicians
* Roberto to send presentation deck and data recruitment package to Al and Harold
* Zheng to visit ALA in person to learn about roadblocks and assess the cost of adding resources
* OSIC creating infrastructure for working groups
* OSIC creating legal entry

**Notes:**

* ALA has 17 sites with 25 institutions
* We would need to have direct negotiations with the smaller organizations, as they have their own legal and technical requirements.
* ALA can help leverage relationships
* The two assets OSIC needs from ALA is expertise and data
* Need to discover which data ALA has the rights to at each institution
* Concern is how far up the ladder we would need to go to get the approval to share the scans
* Data compilation stage is the trickiest. OSIC will work with each institution during this phase. Can provide monetary resources to help get this done. Once compiled, the transition is easy
* The Data working group will decide what data needs to be sent
* Website will be launched until after legal entity is established

**OSIC meeting with the University of Iowa   
Attendees:** Roberto, Eric, Amin, Joe, Megan

**Action Items:**

* University of Iowa to provide rep for Data WG to Megan and Roberto
* OSIC creating infrastructure for working groups over next two weeks
* OSIC creating legal entity in one month
* OSIC creating Data WG in next two weeks

**Notes:**

* Emil – engineer, Alicia – physician
* 300 scans to share
* OSIC has long and short-term goals, and commitment from its founding partners for this 4-year project.
* Legal framework should enable technical decisions by the members on how to operate.
* Data WG will work on minimal atomic data required, data collection, compilation, and anonymization in the next 2 months
* High special frequency algorithms may present a problem. To be discussed at the Data WG.
* Image and data will be included in an atomic pack. No way to link to a future scan of the same statement
* May need to anonymize the data before it leaves Iowa – discuss with IRB.
* The institution can always re-identify the data, but no one else can.
* Any data collected in the future we can get consent. Legacy data needs this system. Consent can come from family if patient is deceased, depending on geographic location.
* Data WG will be discussing IPR
* Data WG will create data set requirements and access controls
  + Option 1: only data contributors will have access to the database
  + Option 2: all OSIC members will have access
* New data for each challenge- data quality may vary. Therefore we are building standards
* Divide into training, testing, validation sets – keep the test set locked that will be used by judge
* Make sure test set is not biased toward institution that contributed. To solve this, randomize from a large selection of institutions
* Data governance model to be discussed in the Data WG
* Consider training time, if the database is very large. Look for donations of processing power. (get from Amazon)
* PFF foundation are organizing a database. Need to figure out if data will be collected directly or through PFF
* Concern that the provider could fingerprint data? Need to discuss in WG. The reward is personal money, so there is a financial incentive for misbehavior.
* Data WG will decide what data points are needed
* Could make rule that institution is not given access to own data
* If it is too complicated or two risky, it will be difficult for some institutions to join. Large brands should help bring in smaller institutions.
* The institutions will identify what resources they need to compile the data, and OSIC can mitigate some of the costs. No funding for the institutions to hire post-docs to develop algorithms. OSIC is funding David Barber’s team at UCL. OSIC will not be creating a foundation, and would need more funding
* Official launch will be at ERS, including Kaggle challenge