Landscape analysis for a disease progression model of bronchopulmonary dysplasia (BPD) in neonates

Leveraging clinical trial experience and real-world patient data

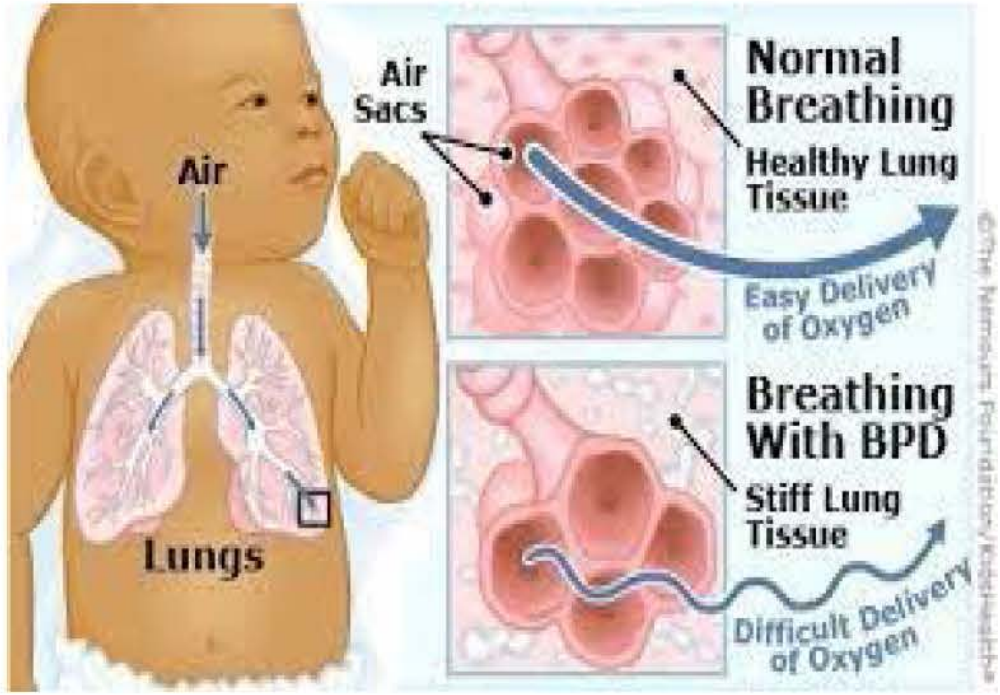
INC INTERNATIONAL NEONATAL CONSORTIUM

CRITICAL PATH INSTITUTE 15 YEARS

Grant Opportunity – Regulatory Interest

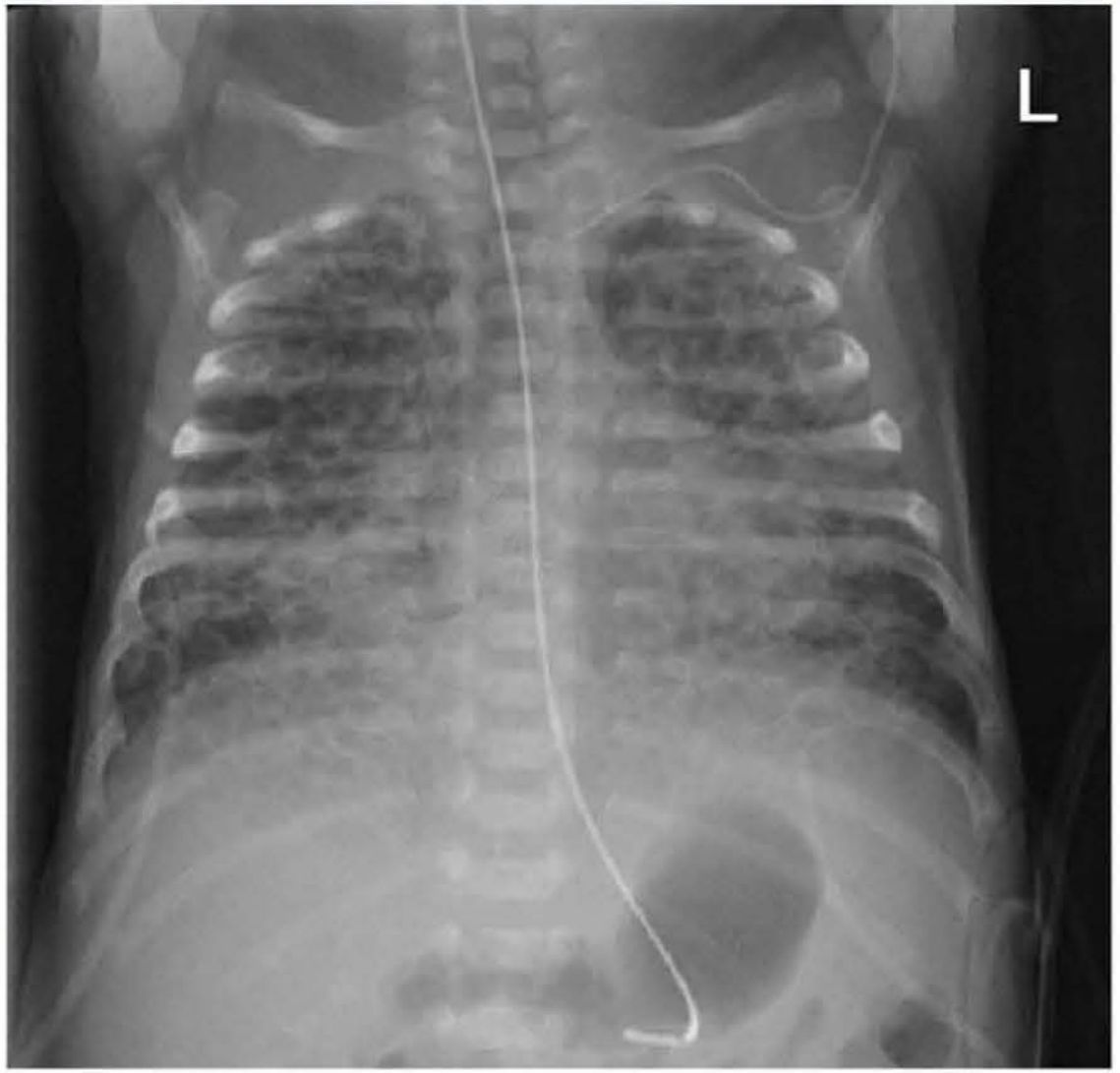
- Using Real-World Data (RWD) to generate Real World Evidence (RWE) in Neonatology
  - There is an urgent need to identify risk factors to enrich the population being studied and develop a more **comprehensive disease progression model of Bronchopulmonary dysplasia (BPD).**
  - There is a critical need to **define laboratory value reference ranges** for the neonatal population.
- C-Path will develop a Real-World Data and Analytics Platform (RW-DAP)
  - Data will be generated from EMRs, Clinical Trials, and other data sources to:
    - Develop a quantitative disease progression model for BPD
    - Generate reference ranges for all commonly used laboratory values
    - Identify gaps that interfere with the optimal use of existing RWD
    - Identify solutions that improve the usability of RWD in future

About BPD - Bronchopulmonary dysplasia



About BPD

Bronchopulmonary Dysplasia



Prevention

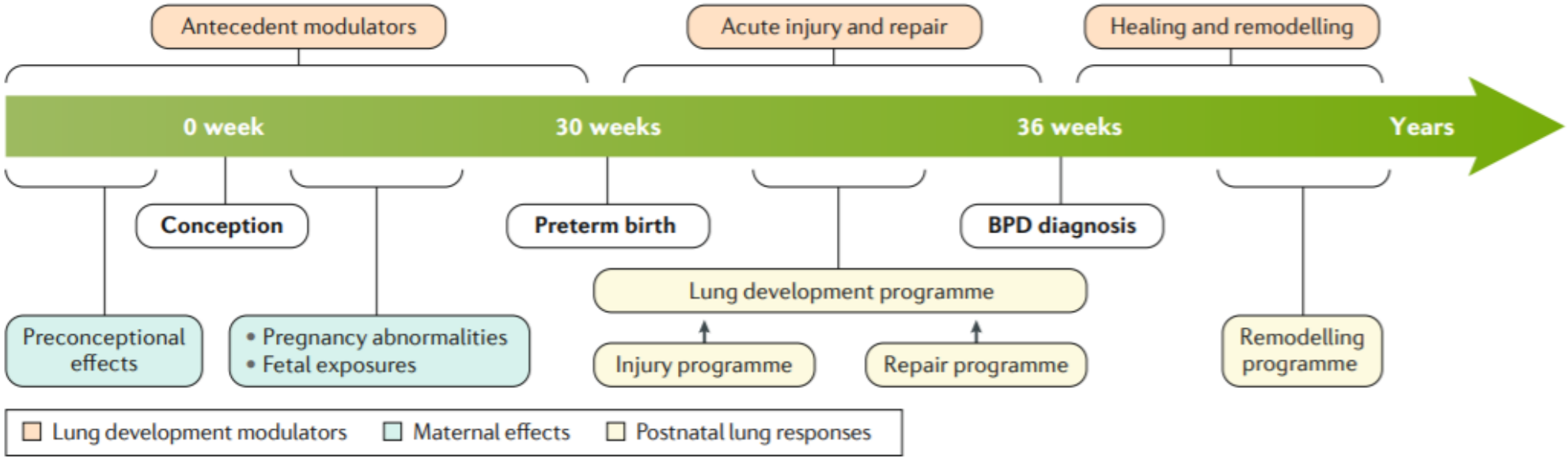
- Antenatal corticosteroids
- Surfactant replacement therapy
- Caffeine
- Ventilation strategies
- Postnatal corticosteroids
- Macrolide treatment of Ureaplasma spp.
- Fluid and nutrition management
- Vitamin A

Therapy

- Fluid and nutrition management
- Diuretics
- Postnatal corticosteroids
- Stem cell therapy
- Therapy of BPD associated pulmonary arterial hypertension

BPD Disease Progression

- The timeline indicates variables that may modulate lung development from preconception through fetal development before preterm birth. Acute injury (on the timescale of days and weeks) resulting from neonatal care that is required to ensure survival then progresses to chronic lung injury and, ultimately, repair and remodeling over months and years.
- The unique aspect of bronchopulmonary dysplasia (BPD) is that it is an injury process that occurs as the premature lung is being injured and must repair as the lung continues to develop and mature. Remodeling of the lungs can occur over years.



Where does the data come from?
INC Membership

- Children's Hospital Association
- Industry Clinical Trials
  - Bayer
  - Chiesi
  - Novartis
  - Takeda
  - Others
- Academic Clinical Trials
- Hospital EMRs
  - I-ACT network (>120,000 neonates)
  - Others
- Japanese Neonatal Network
- Others

Approach to Data Management

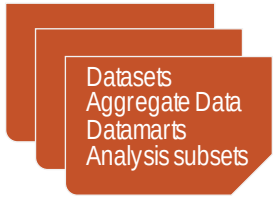
Data Contribution Agreement for each dataset

Curate and Standardize Data

Data Extraction

- Cloud based portal with controlled access to approved users
- Multi-factor authentication
- Tiered access to data based on scope of disclosure permitted in DCA
- Authorized users can export aggregate DB, select files, datamarts, custom reports, etc.

Aggregate/Store Data in DB



Extract Data and Analysis Datasets

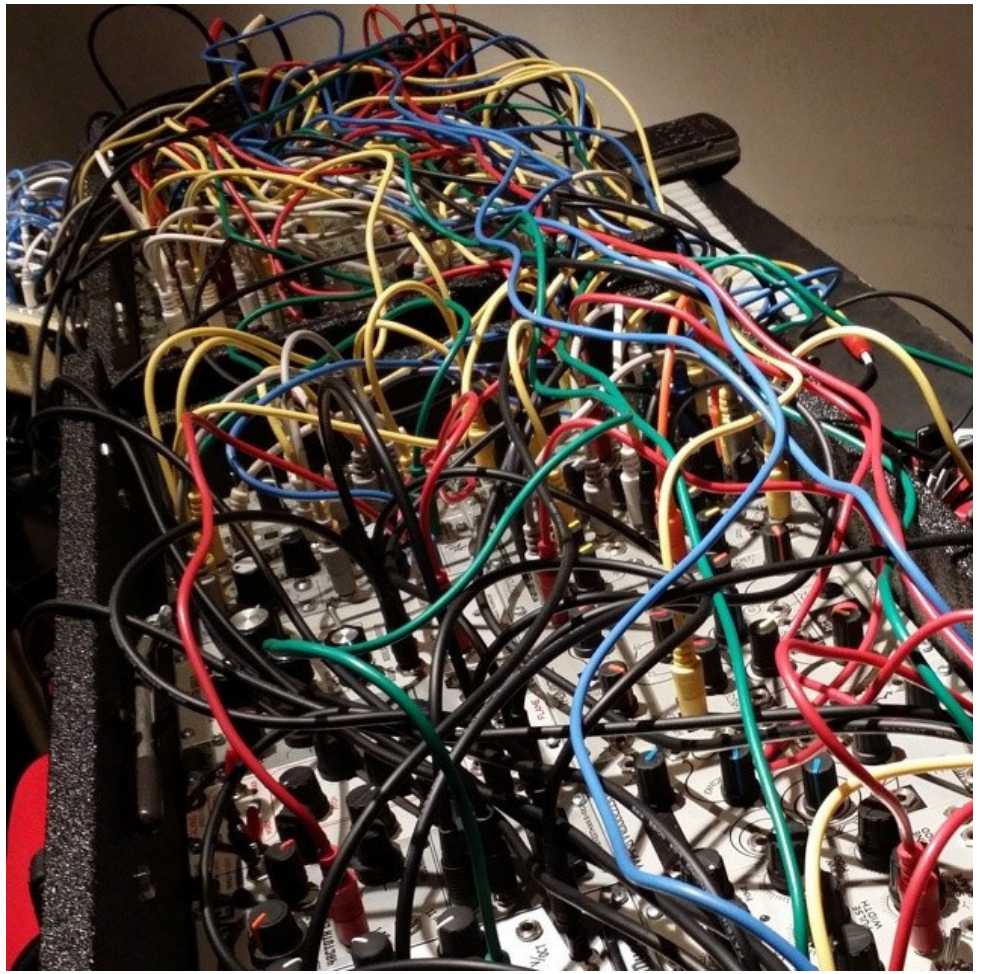
Real World Data and Evidence

INC
INTERNATIONAL NEONATAL CONSORTIUM
CRITICAL PATH INSTITUTE

- Challenges Associated with RWD and RWE
  - Identifying Sources of Data
  - Patient Privacy and Data Security
  - Lack of Harmonized Data Structures
  - Varying Data Quality

Data Harmonization and Integration

- Data from multiple sources is like a big mess of tangled cables
  - Cables may be for power, video, audio, or other functions
  - Multiple plug types mean they can't be interchanged easily
  - Every "data source" is a unique plug type



Data Harmonization and Integration

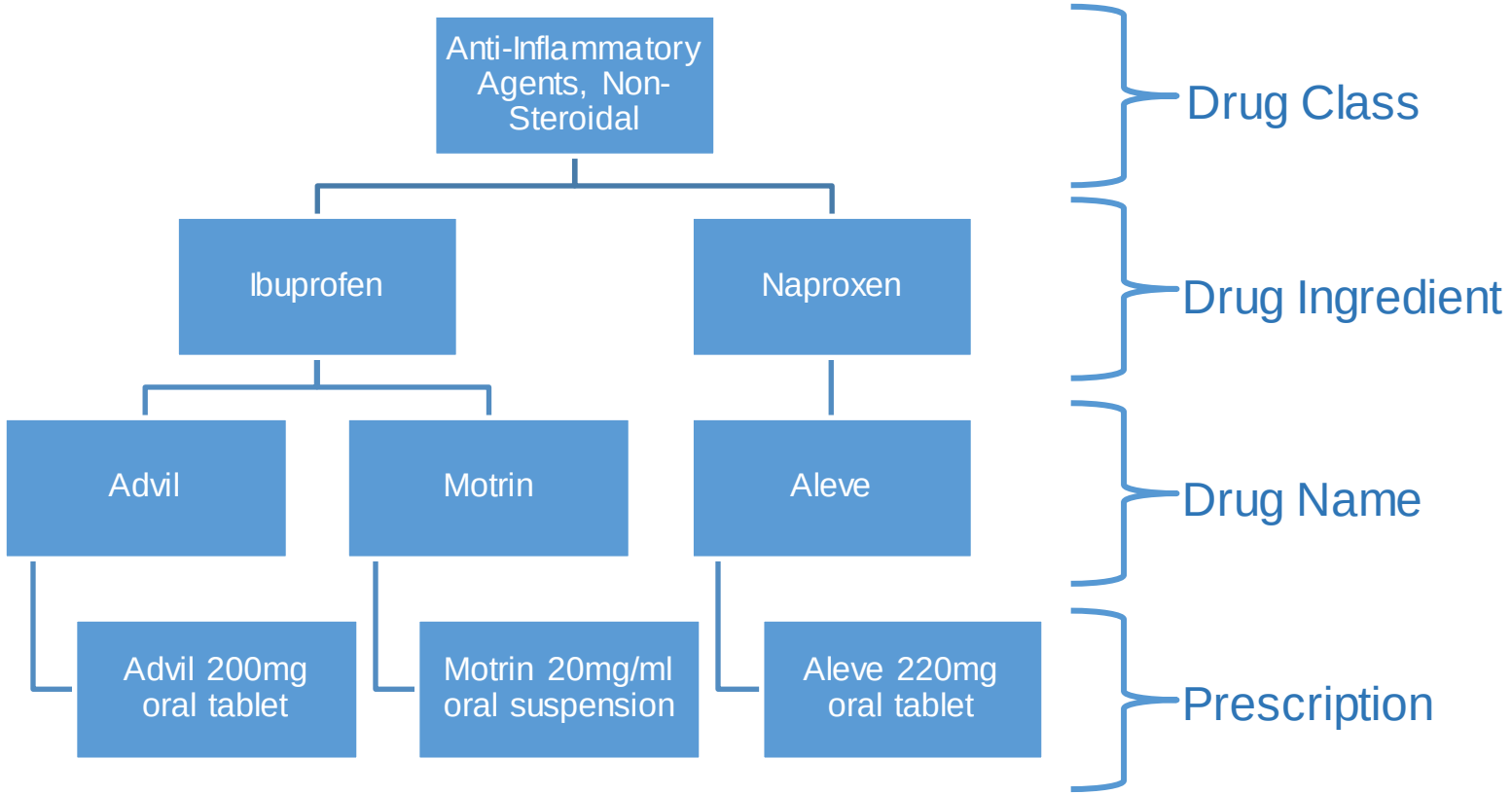
- Curation and harmonization transforms “plugs” into a single standard plug type
- Mapping to a single standard requires harmonization of semantic meanings
- Data paradigms are diverse but can be accommodated
  - Metadata tells us how to interpret a plugs purpose
    - “Is this power or audio?”
  - Some “plugs” already have “adapters” to convert data
  - Other plugs need manual mapping

Common Data Model


Manual Mapping

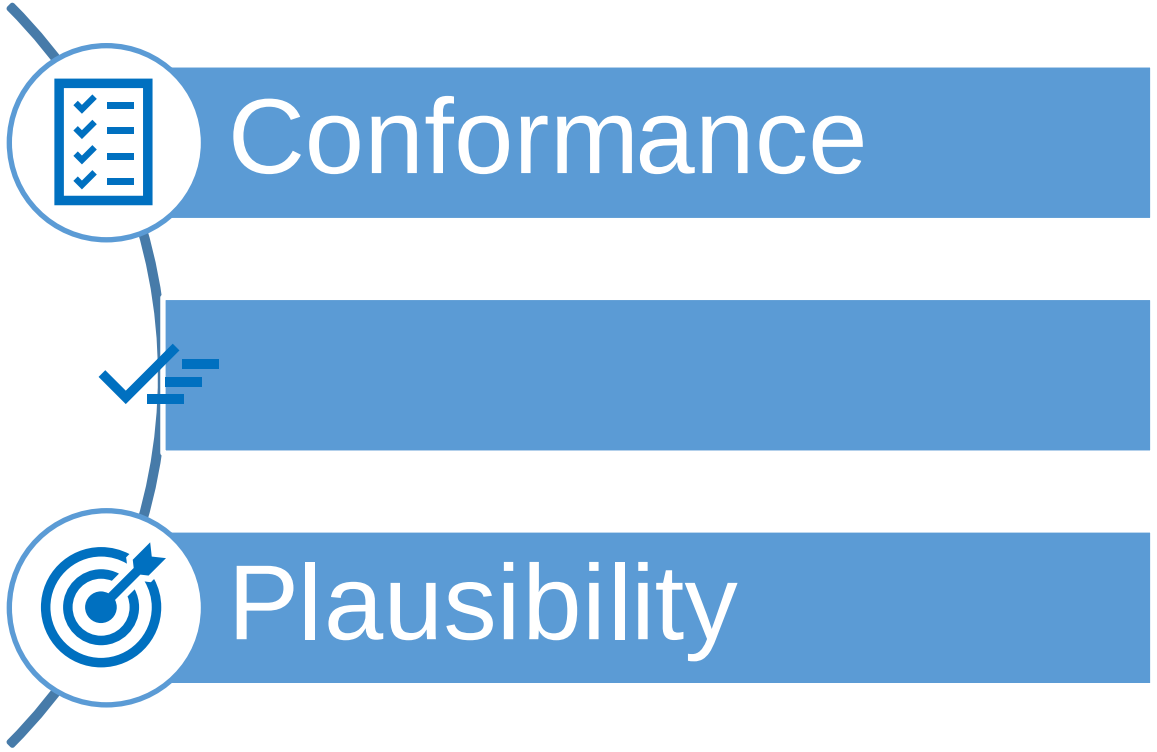
Data Harmonization and Integration

- Data is harmonized into a standardized vocabulary with a hierarchy of related concepts
  - Allows for data of varying “granularity” to be integrated into a single source
  - Structures data around clinically relevant concepts in a parent-child fashion
  - Can get very complicated with new concepts and multiple concepts!



Evaluate Data Quality

- Examples of Data Verification and Validation:
  - Conformance
    - *Does the data comply with the standards?*
  - Completeness
    - *What are the frequencies of key data elements?*
  - Plausibility
    - *Are the data values believable and reasonable?*



Data Quality Approach

C-Path plans to implement multiple validation and verification checks including:

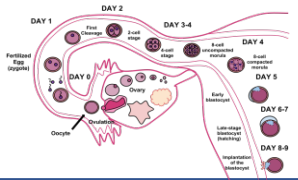
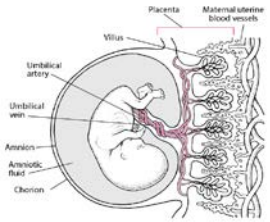
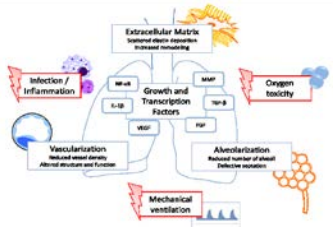
1. Conceptual and Operational Definitions
    1. In partnership with INC working groups C-Path will develop conceptual definitions of key study variables such as conditions and outcomes of interest that can be operationalized within the context of the data
    2. Operational definitions will be developed against the conceptual definitions to define the outcomes of interest within the context of the integrated dataset
    3. Study variables will be compared to available benchmarks, when possible, to estimate any potential misclassification and sources of bias within the data sources
2. Evaluate data quality metrics using guidelines and checklists for EHR and data warehouses[i]
    1. Data quality checks will occur during data harmonization processes, after harmonizing to a standard model, and after integrating data sources into an analytical subset
    2. Utilizing open-source tools such as the OHDSI Data Quality Dashboard[ii] on the transformed data in the CDM to evaluate 20 check types with over 3,000 individual data quality checks to pass rates of individual and integrated datasets
    3. Applying additional context relevant data quality checks to evaluate the conformance, completeness, and plausibility relevant to the disease and population
3. Conduct regular QA and QC Steps including:
    1. Tracking and documenting the provenance of core data elements throughout each data processing step
    2. Applying consistent handling of data discrepancies and errors with clear documentation and traceability
    3. Reviewing accuracy of data mappings that maintain semantic equivalence within- and- between different coding systems

i. Miksad and Abernethy 2018; Girman et al. 2018; Daniel et al. 2018; Kahn et al. 2016; Wang et al. 2017; Mahendraratnam et al. 2019
ii. https://ohdsi.github.io/DataQualityDashboard/index.html

BPD Project Cohort Specification

Observation Period	January 01, 2010 – December 31, 2020
Entry Event	NICU admission during 22 - 42 weeks gestational age
Follow Up Period	Longitudinal follow up through childhood
Associated Persons	Include data for neonate and for mother

What do we measure? What can we measure? What should we measure?

Timeline	Development Phase
–	Conception
Week 0	
Week 30	Preterm birth
35-40 Weeks	 BPD Diagnosis
. . . . Years	

Clinical Events

Preconceptional effects
- Genetic susceptibility
- Antenatal factors: chorioamnionitis

Postnatal Factors
- Ventilator traumas
- Oxygen toxicity
- Pulmonary edema
- Sepsis
- Pregnancy abnormalities
- Fetal exposures

Injury program
- Diuretics
- Bronchodilators
- Corticosteroids
- Viral immunization
- Cardiac medications

Repair program
- Impaired alveolarization
- Dysregulated angiogenesis
- BPD

Remodeling program

Measured Lab Tests*

Neonates (up to 44 weeks postmenstrual age)
- Crs (passive respiratory system compliance), Rrs (passive respiratory system resistance), Cdyn (dynamic compliance), FRC (functional residual capacity), Z_{rs} (Impedance of respiratory system), and LCI (lung clearance index / mixing efficiency)

Infants (Age 44 weeks PMA to 2 years of age)
- As above for neonates, Crs, Rrs, Cdyn, FRC, Z_{rs}, and LCI
- FEFs (forced expiratory flows), FVC (forced vital capacity), FRCp (functional residual capacity), DLCO (pulmonary diffusing capacity) and VA (alveolar volume)

Children / Adolescents / Adults
- FEFs, FVC, FEV1 (forced expiratory volume in 1 second), FRC, DLCO

*Pulmonary function tests used to monitor the natural history of BPD (McEvoy et al, 2016)

Requested Data Domains

INC
INTERNATIONAL NEONATAL CONSORTIUM
CRITICAL PATH INSTITUTE

Neonate Characteristics

- Demographics
- Anthropometric measurements
- Comorbidities

Laboratory Tests and Measurements

- Blood Chemistry, Panels, and Urinalysis
- Hematology and Cell Counts
- Vital Signs

Maternal Characteristics

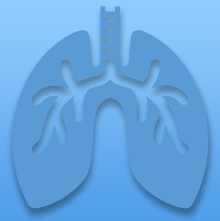
- Demographics
- Medical and Substance Use History
- Antenatal Medications

Respiratory Metrics

- Oxygen Administration and Measurements
- Respiratory Findings
- Procedures and Respiratory Support

Hospitalization and Clinical Encounters

- Admission and Disposition
- Length of Stay and Timing
- Longitudinal Healthcare Utilization



Additional and Disease Specific Metrics

- Medications and Breast Milk Administration
- Brain Imaging Studies
- APGAR, Bayley-3, and Griffin Scores

INC Data Received To Date

Contributor	Type	Study Name	N Participants	Focus Area
Tufts Medical Center	Clinical Trial	CC-10	88	BPD
Tufts Medical Center	Observation Study	STOP-BPD	86	BPD
NICHD-DASH / NRN	Clinical Trial	SUPPORT	1316	BPD
Tufts Medical Center	Clinical Trial	NAS	700	NAS
Bayer	Clinical Trial	EINSTEIN Junior	500	Venous Thrombosis
Bayer	Clinical Trial	FIREFLEYE	113	ROP

Data Contributions In Progress

Contributor	Type	Status
Tufts Medical Center	Electronic Health Records	Executed Agreement, Pending Data Transfer
Chiesi	Clinical Trial	Executed Agreement, Pending Data Transfer
University of Colorado	Observational Study	Executed Agreement, Pending Data Transfer
Japan NRN	EHR Network	Agreement Under Negotiation
CHOP	Electronic Health Records	Agreement Under Negotiation
U of Utah	Electronic Health Records	Agreement Under Negotiation
U of Missouri	Electronic Health Records	Agreement Under Negotiation
U of Minnesota	Electronic Health Records	Agreement Under Negotiation
U of New Mexico	Electronic Health Records	Agreement Under Negotiation

Received Data Inventory

Data Category	INC-1001	INC-1002	INC-1004
Neonatal Demographics	✓	✓	
Maternal Demographics		✓	
Maternal Medical History and Substance Use History		✓	
Laboratory Measurements	✓	✓	
Hospitalization Admission and Disposition		✓	
Oxygen Administration and Respiratory Support Procedures		✓	
APGAR Scores	✓	✓	
Bayley-3 Scores	✓		
Griffin Scores			

✓ = Available
✓ = Partially Data Elements Available

BPD data and analysis

Data to be collected:

1. Gestational age, postnatal age, birth weight, length, head circumference, sex, race/ethnicity, maternal age
2. Medicaid vs. private pay
3. Other potential maternal confounders – smoking, alcohol use, drug use (licit, illicit), obesity, diabetes, hypertension/pre-eclampsia, etc.
4. Presence of chorioamnionitis/early onset neonatal sepsis
5. Oxygen administered, oxygen saturation (%), partial pressure of oxygen (PaO_2), partial pressure of carbon dioxide ($PaCO_2$), pH, bicarbonate, base deficit
6. Requirement for invasive/non-invasive respiratory support
7. Clinical signs of respiratory disease (tachypnea, retractions, rales/wheezing)
8. Chest radiographic findings of persistent densities in both lungs, air trapping, atelectasis
9. Antenatal corticosteroids, magnesium, tocolytics, or other medications in 3rd trimester
10. Breast milk administration/duration
11. Parenteral nutrition administration/duration
12. Significant confounders such patent ductus arteriosus (hemodynamically significant; medical/surgical treatment), sepsis, necrotizing enterocolitis, significant abnormalities in brain imaging, and early mortality that can impact BPD and longer-term outcomes.

Long-term outcomes that correlate with clinically meaningful outcomes:

1. Late death
2. Serious respiratory morbidity (tracheostomy, hospitalization beyond 50 weeks corrected gestational age, supplemental oxygen/ventilatory support/respiratory monitor use, repeated hospitalizations, use of respiratory medications, non-routine medical visits)
3. Blindness, deafness
4. Bayley-3 cognitive, motor, language composites as well as other neurocognitive scores

Data Analysis:

- Data harmonization
- Multivariate mixed-effect models as a function of SaO_2, PaO_2, $PaCO_2$, pH, etc.
- Validation of pre-specified definitions with diagnosis and long-term outcomes
- Correlation of risk factors with long term morbidity, mortality, adverse respiratory and neurodevelopmental outcomes

Predictive Model Landscape

Model (Reference)	Category	Data Sources	Purpose
Review of 23 clinical prediction models (Onland et al. 2013)	Regression-based models; multivariate analyses; retrospective and prospective	A variety of data sources including the PreVILIG database	• Review the quality and validity of models that predict BPD in preterm infants using clinical information from the first week of life • Attempted to externally validate models and compare predictors identified
BPD severity prediction model (Nino et. al, 2020)	Binary logistic regression models to evaluate the predictive value of different variables, using respiratory hospitalization as the primary outcome	• Primary cohort included 188 premature infants (≤32 weeks PMA) admitted to the NICU at Children's National Health System (CNHS) in Washington, D.C. • The validation cohort included 130 premature infants (≤36 weeks PMA) admitted to the NICU at The Hospital Militar Central and the Hospital Universitario Clinica San Rafael in Bogota, Colombia.	• Approach improved BPD risk assessment, particularly in extremely premature infants. • Internal validation included lung X-ray imaging and phenotypical characterization of BPD severity levels. • External validation conducted in an independent longitudinal cohort of premature infants (≤36 weeks PMA, n = 130; Bogota).
BPD risk prediction model (Alvarez-Fuente et. al, 2019)	Multivariate logistic regression model to identify risk factors for BPD development by determining the odds ratio of both groups, no-BPD versus BPD, in relation to clinical, echocardiographic and analytic factors	• 5 Spanish hospitals: 50 patients with a median gestational age of 26 weeks and weight of 871 g (range 590-1200g).	• Study and model aimed to explore the ability of clinical, echocardiographic and analytical variables to predict moderate or severe BPD in a cohort of extremely preterm infants.
BPD severity prediction (Valenzuela-Stutman, 2019)	Forward logistic regression models with predictive values evaluated using a ROC curve	Multicenter study including 16,407 infants weighing 500-1500 g (2001-2015) from the Neocosur Network	• Predictive power models for moderate/severe BPD and BPD/death at four postnatal ages. • Birth weight contributed the most in explaining BPD, followed by GA and 1-min Apgar score

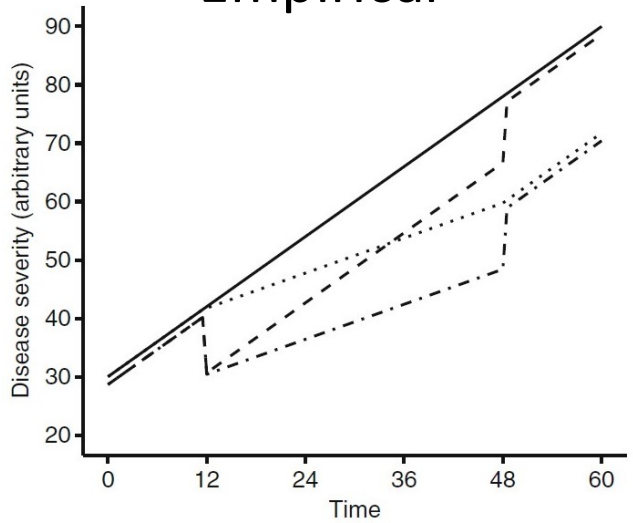
Predictive Model Landscape

Model (Reference)	Category	Data Sources	Purpose
BPD Risk factors in preterm infants (Ding et. al., 2020)	Multiple logistic regression analysis: sensitivity and specificity of the model assessed by ROC curve	Seventy-two preterm infants (30 with BPD and 42 non-BPD controls) admitted in the NICU of the Children's Hospital of Soochow University during 2017 enrolled; prospective longitudinal study	• To identify postnatal risk factors for bronchopulmonary dysplasia (BPD) development in preterm infants with gestational age ≤32 weeks • Perinatal data, a neonatal critical illness score (NCIS), different soluble B7-H3(sB7-H3), and interleukin-18 (IL-18) levels by days after birth collected; early predictive model for BPD development established
Mechanistic model of gas exchange and ventilation under a broad range of local and systemic inflammatory stimuli (Reynolds et. al., 2010)	Diffusion of oxygen and carbon dioxide, hemoglobin uptake of oxygen, and enzymatic reactions governing carbon dioxide and bicarbonate levels.	ODE-based multi-scale model based on literature priors	• Simulation model of pulmonary function under inflammatory stress and of interventions aimed at improving gas exchange in this broadly relevant context. • Generically multiscale model to be improved in its physiologic accuracy and computational load.
Integrative anatomically-based model - incorporates descriptions of material properties and anatomical structure at a range of levels of interest (Tawhai et., al., 2003)	Finite element meshes of the lung lobes, airways, blood vessels, parenchyma, and microcirculation.	Database of publications, models, and data related to the pulmonary circulation	• Provides a framework for quantitative description of a system's geometry, behavior, or interactions. Laboratory observations are incorporated into the models as, for example, geometric data or rate constants, and experimentation is also used to validate the model's performance. • "Lung Atlas" is currently being developed based on structural and functional computed tomography (CT) imaging (Li et al., 2003)

Disease progression modeling approaches

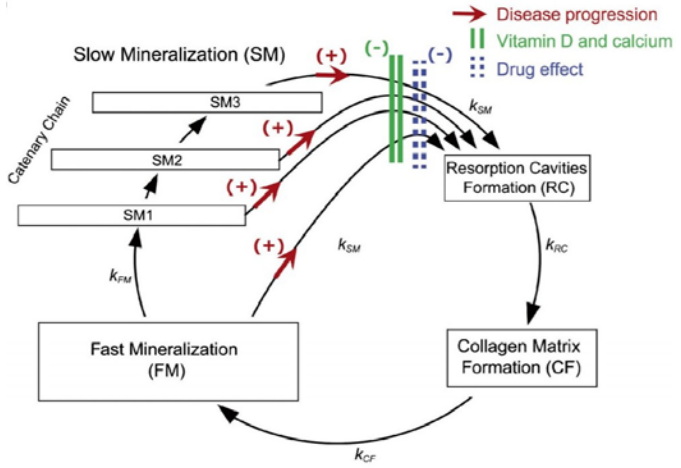
- Using mathematical functions to *quantify* disease dynamics over time
- Models are categorized into three classes:

Empirical



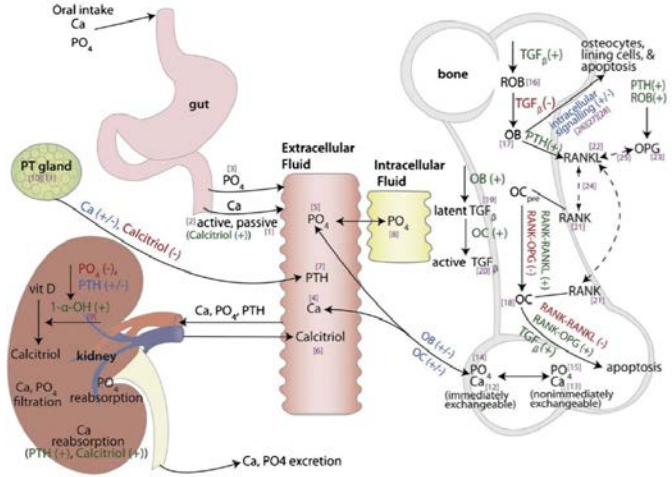
Purely data-driven & what we'll discuss today

Semi-mechanistic



Between empirical & systems biology models

Systems Biology

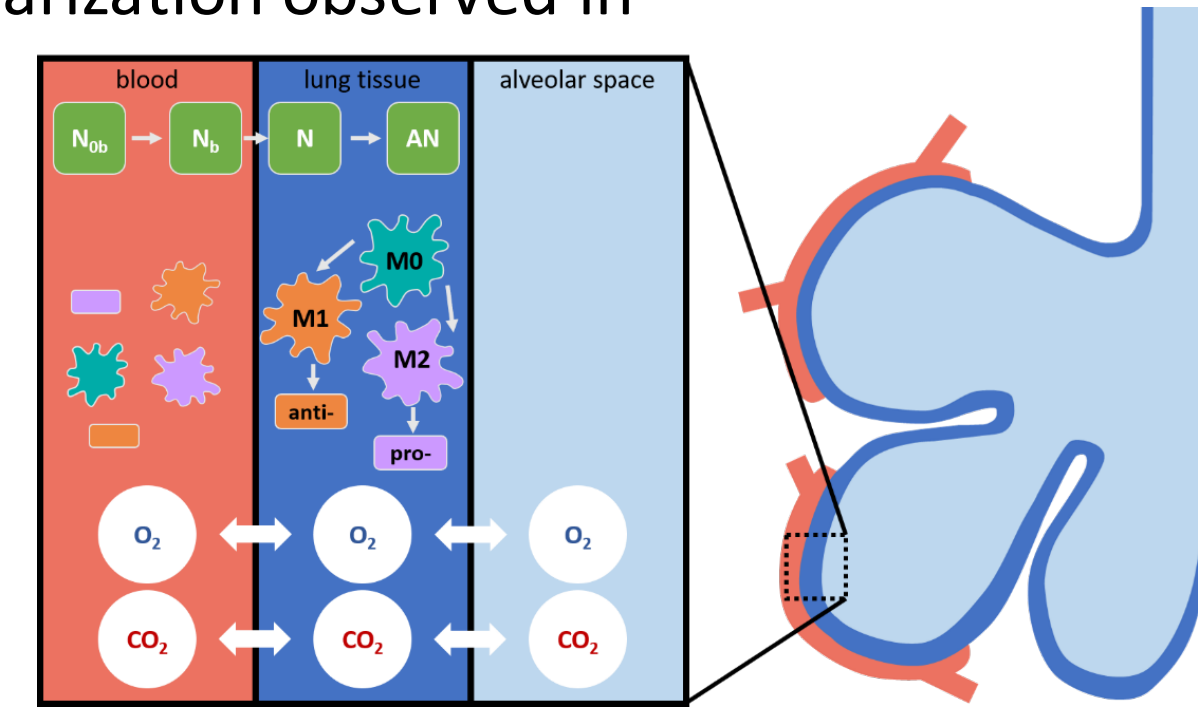


Physiologically-based & incorporate underlying biology

Cook and Bies 2016

Proposed model of lung function dynamics

- Dynamics of gas exchange and key inflammatory mediators can be described with a system of differential equations
- Investigate link between the pro- and anti-inflammatory imbalance and the emergent phenotype of alveolar simplification and dysregulated vascularization observed in BPD
- Potential to link such a model to clinical outcomes like pulmonary function tests or a discrete outcome like the probability of BPD



Modified from Minucci 2021 and Reynolds 2010

Model Framework: Connecting RWD to the model

Proposed Workflow

- Develop mechanistic model for BPD disease progression
  - Borrow from pulmonary lung function models and preclinical BPD experiments and data
  - Link biomarkers of interest to measured RWD elements – assess correlation and consider model-based linkages

- Clinical verification / qualification with INC clinical SMEs

- Re-assess model definitions / revise accordingly

- Validate against external RWD sources

Conclusions:

- Ingestion of RWD sources to support the INC grant is proceeding at a good pace
- BPD data accumulated thus far includes both completed clinical trial data, HER data and registry data
- Curation and standardization of all data sources is ongoing with special attention to quality metrics to qualify COU for modeling purposes
- Early modeling strategies to support quantification of BPD modeling strategies include a family of models including mechanistic QSP models in addition to more traditional predictive modeling efforts informed by RWD
- Underpinning early thoughts on the BPD mechanistic approach will utilize pulmonary function mechanics and requisite biomarkers as an anchor to be joined and vetted against traditional clinical measures obtained from RWD sources
- Results will be transparently shared as the work progresses through INC reporting mechanisms and FDA

Acknowledgements:

- Megan Cala
- Will Roddy
- Jack Beusmans
- Marc Gastonguay
- Kanwaljit Singh
- Mark Turner
- Jon Davis
- Klaus Romeo
- INC
- FDA

Thank you!

INC
INTERNATIONAL NEONATAL CONSORTIUM

CRITICAL PATH INSTITUTE
15 YEARS