Learning Maps in Drug Discovery and Development

Andreas Sashegyi, PhD
Eli Lilly and Company

Introduction / Background

- Learning Map: Formal representation of the ***accumulated kowledge*** and as yet ***untested assumptions***, regarding the ***predictability*** of various tests run during lead optimization on indices that inform the question: «What is the potential for this compound to be a viable, valued drug?»
- Quantitative outcome: a ***confidence evaluation***
- Natural framework for ***accommodating uncertainty*** in parameters and assumtions

Introduction / Background

Objectives of the Learning Map Approach

- Facilitate teams going through a clarification of their ***decision processes***, identify eventual roadblocks
- Provide an efficient way to summarize the conclusions of team discussions and thus to ***share information*** with non-team members (e.g. governance committees)
- Provide and use a documented, transparent, *a priori* defined, ***quantitative-based decision process***
- Compare compounds within a target [project level] and across targets [portfolio level] based on a ***formal confidence evaluation***
- Help to understand the overall ***structure of the project***

Learning Map Key Concepts

- Try to focus on the end point – getting the drug on the market – not just to the clinic
- Not process-oriented (like a process map) but prediction-oriented
- "Why" (Learning Map) versus "When" (Process Map)
- Try to be as comprehensive as possible with all the dimensions even if some of them are not going to be tested – Not knowing/uncertainty should impact the confidence evaluation!

Learning Map Key Concepts
Building Blocks

- Indices**
- Tests**
- Data**
- Value functions**
  - $f(\mathbf{R}) \Rightarrow [0,1]$
- Calibration factors* (complete information threshold)
- Combination functions
- Weights**
- Value-of-information factors**

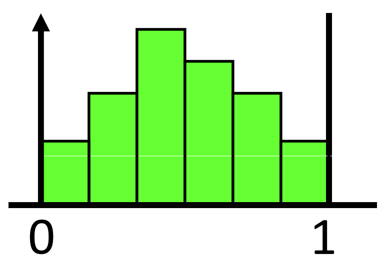
** Team input indispensable
* Team input important

Learning Map Key Concepts

Index versus Test

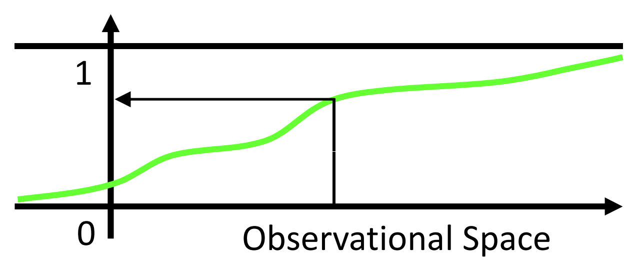
INDEX
A characteristic of a treatment that must be evaluated to determine whether it's approvable and provides meaningful patient outcomes

[0;1]-Histogram



TEST
Represents a unit of information used to assess our confidence about a particular index

Data Transformation into [0;1]



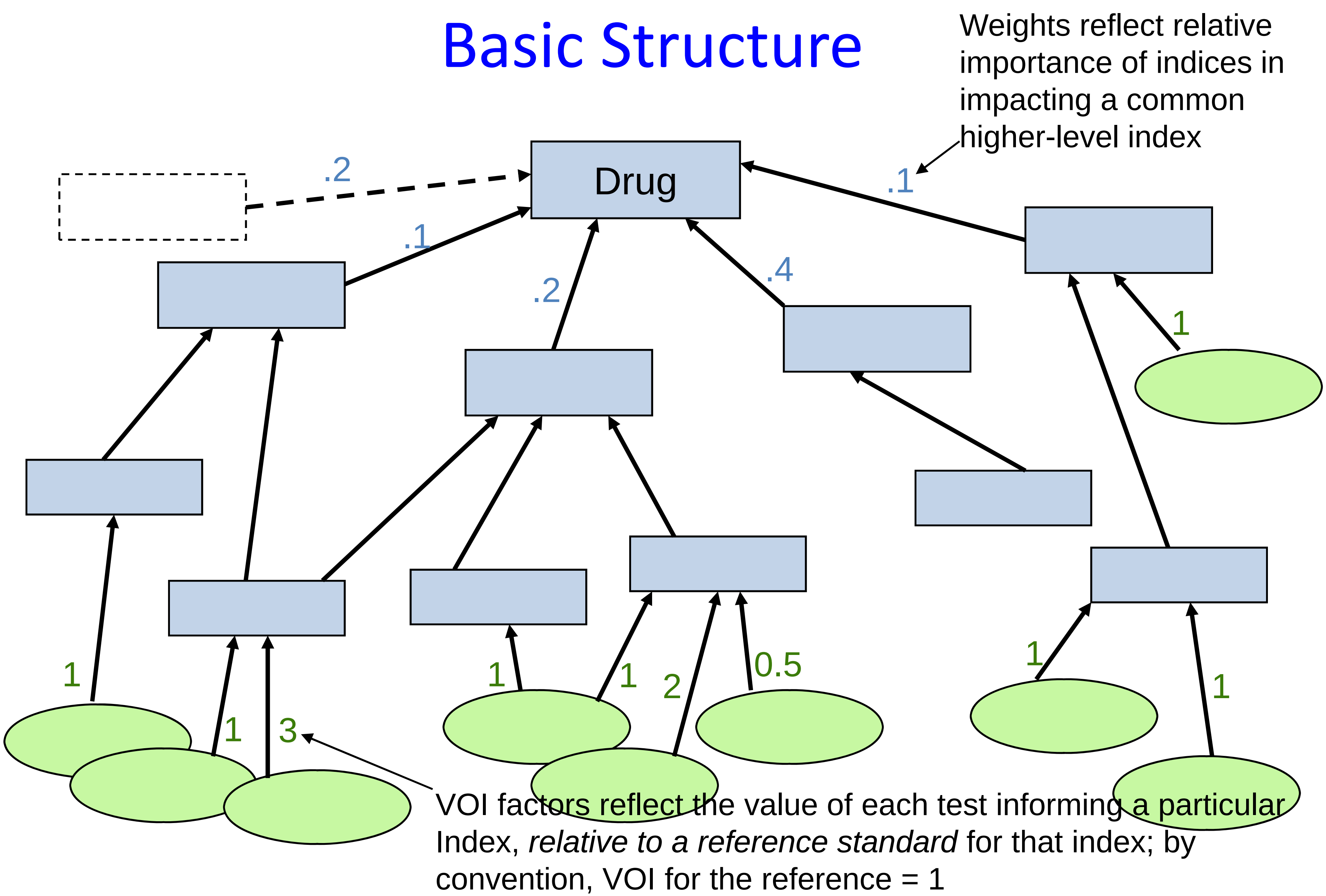
Weighting Factor
Strictly Positive Number

Value-of-info (VOI) Factor
Strictly Positive Number

Data Point
Test Dependent

Learning Map Key Concepts

Basic Structure

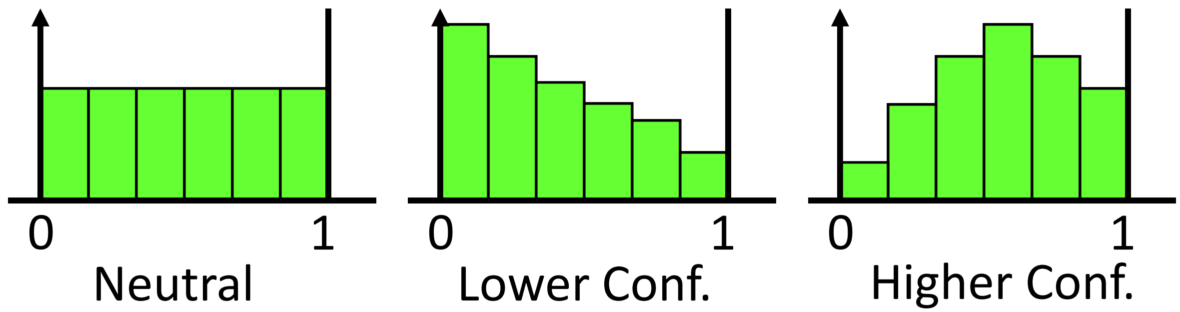


Learning Map Key Concepts

Confidence Quantification

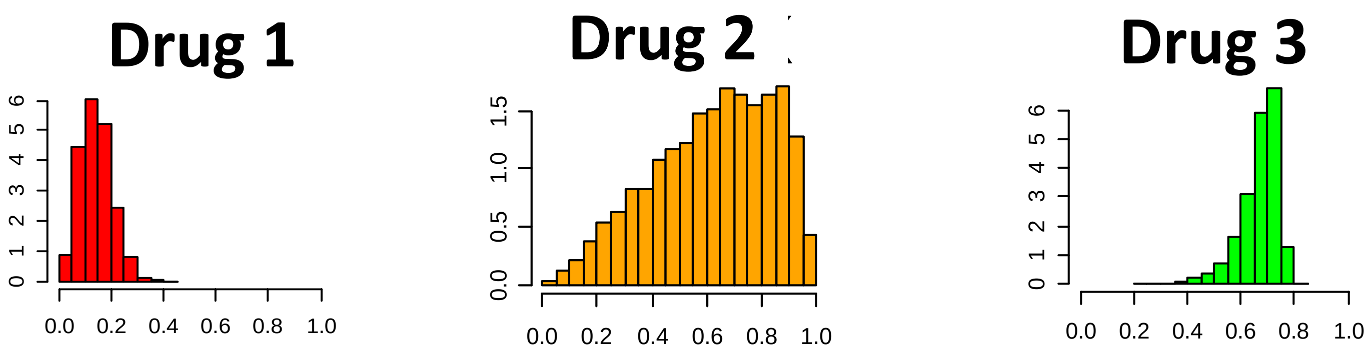
(Confidence Distributions)

Qualitative Inspection



Quantitative Evaluation

Can be summarized / Categorized



Learning Map Key Concepts

Confidence Quantification

- In the absence of information, confidence distributions are flat (uniform on [0,1])
- As data are collected, confidence distributions are updated using a Bayesian approach
- The distributions of all sub-indices that inform a common higher-level index are combined using a weighted average (based on the assigned weights) to yield the distribution of the higher-level index
- Calibration factors for each index regulate how quickly data from the various tests overcome the prior distribution of the index

Some Technical Details

Implementation Based on Beta-Binomial Model

A heuristic approach…

- Assume that a team has defined its LM structure –
  - All indices defined
  - Connections between indices
  - Weights associated with each set of indices impacting a common higher-level index
  - All tests defined, plus transformation functions and VOI factors

- One may use a generalization of the beta-binomial model as an intuitive way of evaluating the LM, that avoids computationally intensive methods

Some Technical Details

Implementation Based on Beta-Binomial Model

Recall the beta-binomial formulation:

$$\text{If } X \mid p \sim Bin(n, p)$$
$$\text{and } p \sim Beta(a, b),$$
$$\text{then } p \mid X \sim Beta(a + x, b + n - x)$$

- Key to this approach: think of a given test result (after transformation to [0,1]) as an observation from a binomial distribution $X|p$ in a beta-binomial model, where
  - p is the prior distribution of the index in question
  - $p|X$ is the posterior distribution of the index, accounting for the impact of the observation

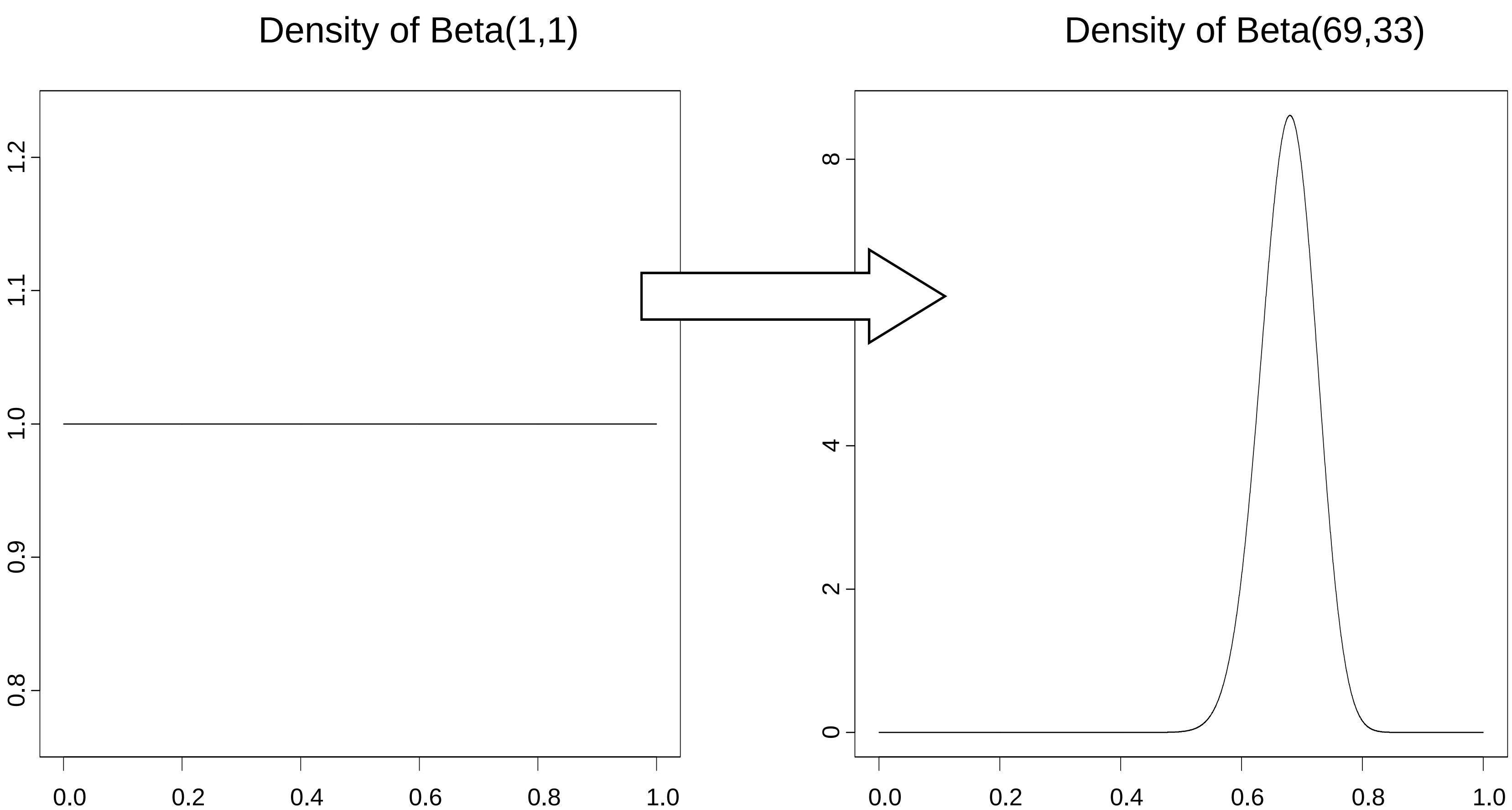
Some Technical Details

Implementation Based on Beta-Binomial Model

- For each index, settle on a suitable *reference unit of information* – perhaps the result of a standard test
- Impact of all other tests that inform the same index is defined in relation to the reference
  – Example: Let the reference be represented by a binomial with n=100 (a test twice as valuable would be represented by a binomial with n=200; a test half as valuable by a binomial with n=50)
    - Suppose the reference test result yields a result of 0.68 after transformation
    - Represent this value on a scale from 0 to n, with n being the best possible result (giving the greatest confidence)
    - Hence with a uniform prior (a=b=1) and n=100, we have x=68, n-x=32, yielding a posterior distribution *Beta*(69, 33)

Some Technical Details

Implementation Based on Beta-Binomial Model



Some Technical Details

Implementation Based on Beta-Binomial Model

In practice a modification is required to calibrate the rate at which a particular prior distribution is overwhelmed by the observed data

- Replace $p|X$ as defined above by introducing an index-specific calibration factor θ, yielding a posterior distribution that falls between the prior and the proper posterior of the beta-binomial:

$$p \mid X \sim Beta(a + \theta x, b + \theta(n - x))$$

- θ calibrates the impact of the observed data
  - Values close to 0 make it very difficult to overcome the prior
  - Values close to 1 approximate the beta-binomial posterior

Some Technical Details

Implementation Based on Beta-Binomial Model

On choosing θ

- For each index, elicit from the team a quantification of how much data, relative to the reference unit of information, would constitute "practically complete information"
- Statistically, define "practically complete information" in terms of the magnitude of the standard deviation of the confidence distribution
  - E.g.: Choose θ to achieve a std dev no greater than 0.05 at a mean of 0.5, the point at which (given constant information) the variance of the beta is maximized

Some Technical Details

Implementation Based on Beta-Binomial Model

On choosing θ (cont'd)

- Example (cont'd): Suppose the team agrees that for a particular index, information equivalent to 8 times the reference unit would constitute "practically complete information"
  - Since reference unit is defined by Bin(100, *p*), complete information can be represented by Bin(800, *p*)
- To achieve a posterior mean and std dev of 0.5 and 0.05, respectively, solve

$$Var(p \mid X) = \frac{E(p \mid X)(1 - E(p \mid X))}{1 + a + \theta x + b + \theta(n - x)} = \frac{1}{4(3 + 800\theta)} = 0.05^2$$

➔ θ = 97/800 = 0.121

Some Technical Details

Implementation Based on Beta-Binomial Model

An alternative (less heuristic but operationally simpler) approach

- Previously, test results were represented by binomial distributions in which *n* provided a measure of the test's value relative to the reference
- For test *i*,

$$p \mid x_i \sim Beta(a + x_i, b + n_i - x_i)$$

was replaced by

$$p \mid x_i \sim Beta(a + \theta x_i, b + \theta\{n_i - x_i\})$$

which can be re-written

$$p \mid x_i \sim Beta(a + \theta n_i y_i, b + \theta n_i \{1 - y_i\})$$

Measure of importance of test result in contributing to complete information on index in question

Transformed test result (number between 0 and 1)

Some Technical Details

Implementation Based on Beta-Binomial Model

- So, the importance of a particular test *i* in contributing to complete information on a certain index was determined by assessing n_i and an index-specific θ separately
- Consider instead a ***single elicitation*** F_i of the importance of a given test *i*, relative to what constitutes complete information for the index it informs; hence

$$p \mid x_i \sim Beta\,(a + \theta n_i y_i, b + \theta n_i \{1 - y_i\})$$

$$\Longleftrightarrow \quad p \mid x_i \sim Beta\,(a + \omega F_i y_i, b + \omega F_i \{1 - y_i\})$$

- Since F_i already incorporates consideration of the index test *i* informs, ω is not index-specific but a universal parameter for the learning map, depending only on the definition of “complete information”

Some Technical Details

Implementation Based on Beta-Binomial Model

- The posterior distribution for a particular index informed by *k* tests is therefore given by

$$p \mid x_1,\ldots,x_k \sim Beta(a+\omega\sum_{i=1}^{k} F_i y_i, b+\omega\sum_{i=1}^{k} F_i\{1-y_i\})$$

- Solving for ω, assume uniform priors (*a*=*b*=1) and the same definition of "practically complete information" given earlier. It then follows that

$$\omega = 97 \big/ \sum F_i$$

- Finally, adopt the convention that for complete information, $\Sigma F_i = 1$ so that
  - ω = 97,
  - F_i ***represents the proportion of complete information furnished by test i for the index that test informs***

Some Technical Details

Implementation Based on Beta-Binomial Model

On combining indices

- Posterior confidence distributions of a given set of indices that all inform a common higher-level index may be combined by resampling from each of the distributions and using a weighted geometric or arithmetic mean, based on the assigned weights
  - E.g.: Let X_1, X_2, X_3 represent distributions of a set of sub-indices having weights w_1, w_2, w_3 and a common parent index with distribution Y
    - Arithmetic mean: $Y \sim w_1X_1 + w_2X_2 + w_3X_3$
    - Geometric mean: $Y \sim X_1^{w1}X_2^{w2}X_3^{w3}$
  - Approximate empirical distribution of Y with a beta, computing its parameters a and b using the appropriate transformations of the observed mean and variance of the empirical distribution

Some Technical Details

Implementation Based on Beta-Binomial Model

On combining indices – back-transformation

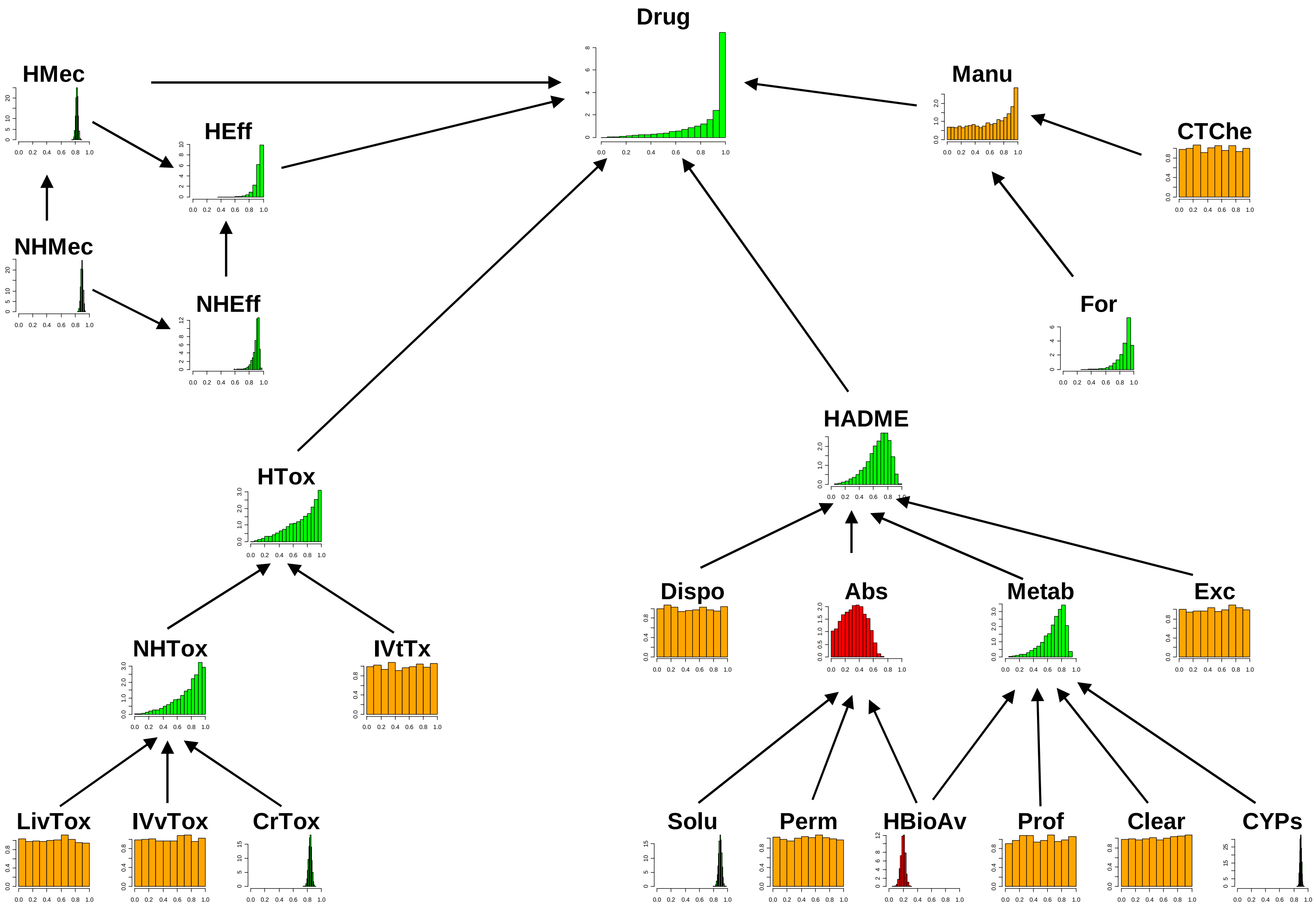
- An important property in interpreting learning maps:
  - If all sub-indices of a common parent index are uniformly distributed, the parent index should also have a uniform distribution

- However, $X_1, X_2, X_3 \sim \text{beta}(1,1) \not\rightarrow Y \sim \text{beta}(1,1)$!

- Having determined $Y \sim \text{beta}(a, b)$, calibrate this distribution by applying a transformation that <u>produces a uniform distribution</u> if each of the sub-indices are uniformly distributed:
  - Suppose $X_1, X_2, X_3 \sim \text{beta}(1,1) \rightarrow Y \sim \text{beta}(a, b)$
  - Let $\delta_1 = 1/a$, $\delta_2 = 1/b$
  - Determine the distribution of the parent index in all cases as $Y \sim \text{beta}(\delta_1 a, \delta_2 b)$, regardless of the distributions of the sub-indices

Sample Visualization

LTB4 Pilot Study Outcomes [Indices]



Objections and Counter-arguments

- A tool can not take decisions for us! ***Best judgment prevails!***
  - Aid your decision-making, not replace it
  - Might highlight some uncovered dimension
- Time consuming to develop a Learning Map – ***Just one more thing we have to do!***
  - Development of Templates/Guidelines/Appropriate Software
- It is ***impossible to be comprehensive***, i.e., to list everything!
  - This is also true when decisions are taken without learning map
- For most of the dimensions, ***decision process is qualitative!***
  - The Learning map approach also holds a qualitative part (the structure of the LM) that can be useful by itself.
- Only a snapshot – ***Science is moving very quickly!***
  - Learning Map should be updated whenever new information is obtained