

INTRODUCTION

The underlying principle of read-across is that biological activity is a function of physical and structural properties of chemicals. Analogs are typically identified on the basis of structural similarity and subsequently evaluated for their use in read-across on the basis of their bioavailability, reactivity and metabolic similarity. While the concept of similarity is the major tenet in grouping chemicals for read-across, a critical consideration is to evaluate if structural differences significantly impact toxicological activity. This is a key source of uncertainty in read-across predictions. We hypothesize that inclusion of toxicokinetic (TK) information will reduce the uncertainty in read-across predictions. TK information can help substantiate whether chemicals within a category have similar ADME properties and, hence, increase the likelihood of exhibiting similar toxicological properties. This current case study is part of a larger study aimed at performing a systematic assessment of the extent to which *in-vitro* TK data can obviate *in-vivo* TK data, while maintaining or increasing scientific confidence in read-across predictions.

OBJECTIVES

1. Perform unsupervised clustering of chemicals with human *in vitro* TK parameters (intrinsic clearance and fraction of chemical unbound to plasma protein) to determine whether structurally related chemicals had similar TK parameters
2. Evaluate the applicability of similarity-weighted read-across predictions of *in vitro* TK parameters for a set of target chemicals using analogs within the clusters identified above

METHODS

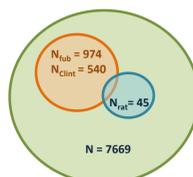
Chemical Descriptor	Description	Similarity Measure
PubChem	881 bits long substructural fingerprint	Jaccard/Tanimoto coefficient
Chemotyper	771 bits long substructural fingerprint	Jaccard/Tanimoto coefficient
PaDel	1875 chemical descriptors (1444 1D, 2D descriptors)	Euclidean distance

Chemical universe

● HTTK exposure data

● HTTK *in vitro* data

● TK *in vivo* data



2. Read-across analysis

Target selection:

Chemicals with rat *in vivo* data (n = 33)

Target cluster assignment:

Assign the target to a KMeans cluster determined in 1.

Analog selection:

- Chemicals with human *in vitro* data
- Similarity threshold range: 0.70 – 0.95
- Analog pool: Chemicals in KMean cluster to which the target is assigned

Read across method:

Similarity-weighted average

Chemical diversity of targets:

Chemical landscape visualization using

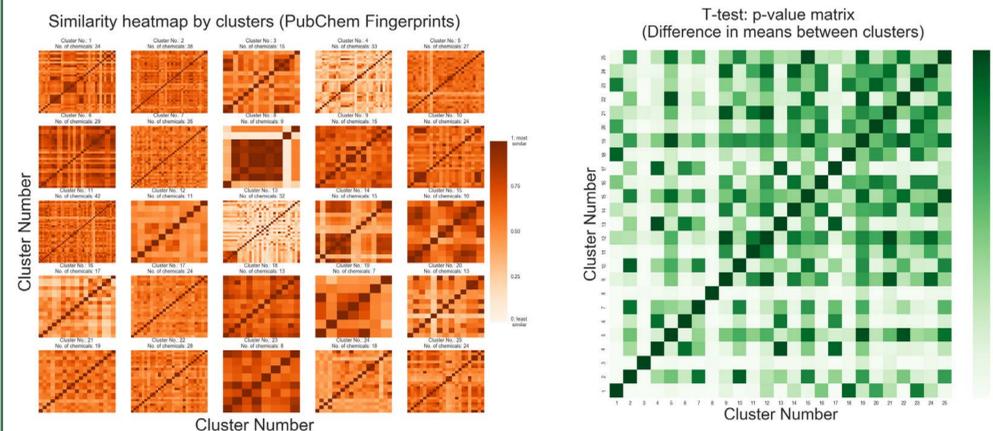
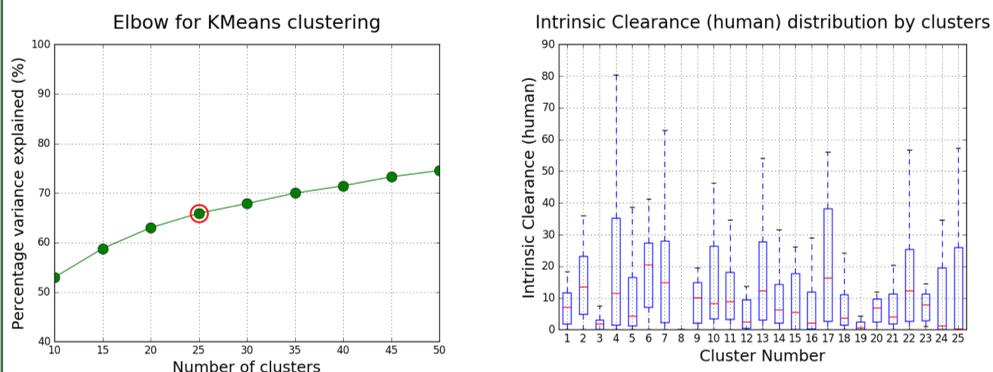
Principal component analysis (PCA)

Number of components = 2

CLUSTERING RESULTS (PubChem)

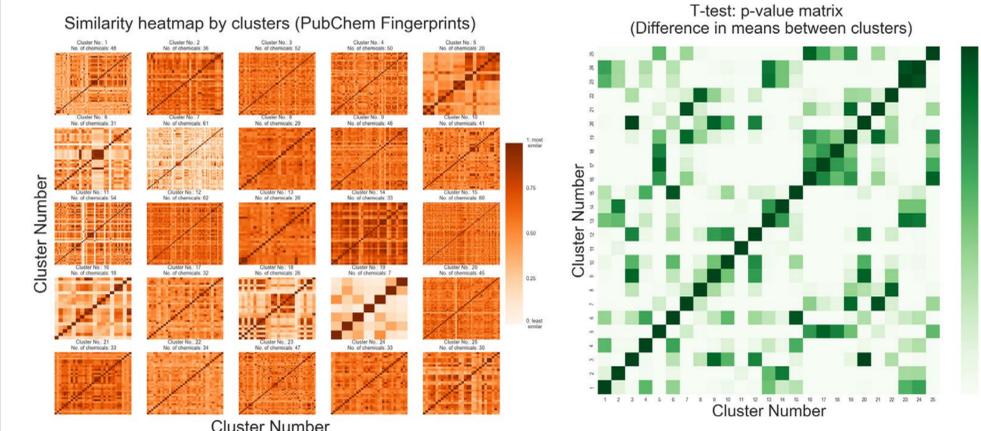
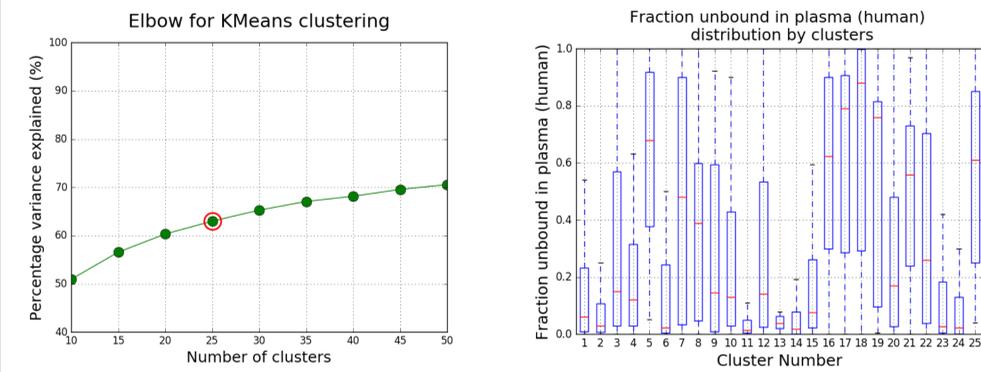
Intrinsic Clearance (Human)

- Feature selection: 159/881
- Optimum cluster number: k = 25



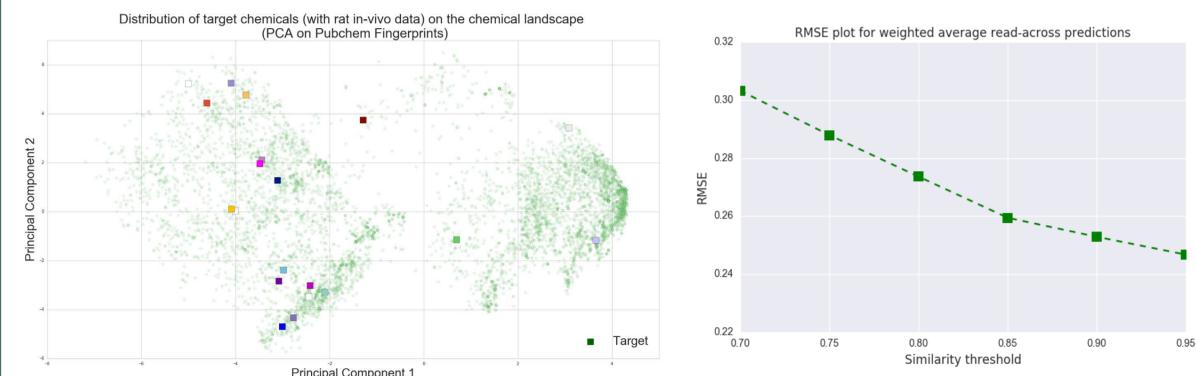
Fraction unbound in plasma (Human)

- Feature selection: 178/881
- Optimum cluster number: k = 25



READ-ACROSS RESULTS

Prediction of fraction of chemical unbound in plasma (*in vitro*) for 20 target chemicals, with rat *in vivo* data, using clusters identified above using PubChem fingerprints



CONCLUSIONS

Unsupervised clustering:

- Preliminary results show a correlation between structural fingerprint-based clusters and TK parameters (intrinsic clearance and fraction of chemical unbound in plasma)
- Adding LogP and PKa as additional descriptors improves the clustering results
- PubChem fingerprints result in more structurally correlated clusters as measured using Jaccard/Tanimoto similarity coefficient

Read-across results:

Prediction metrics in terms of root mean squared errors obtained by weighted average predictions demonstrate promising initial results

Future direction:

Analysis of the correspondence between *in-vitro* TK data and *in-vivo* TK data using the target-analog chemicals identified herein