



Linde Schoenmaker<sup>1</sup>, Willem Jespers<sup>1</sup>, VHP4Safety consortium, Joost Beltman<sup>1</sup>, Gerard van Westen<sup>1</sup>

1. Division of Drug Discovery and Safety, Leiden, The Netherlands



### Improve bioactivity prediction

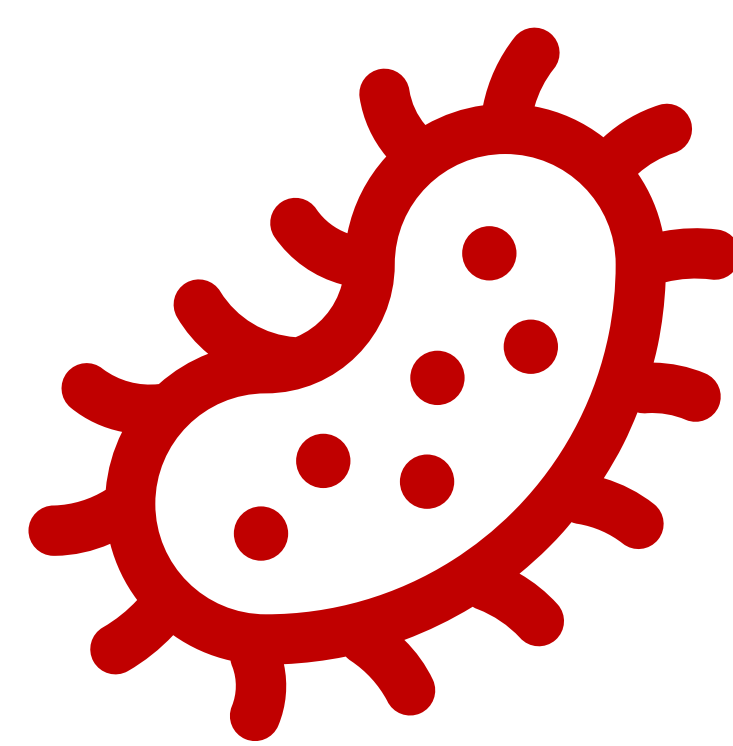
When using **open-source** data for protein-ligand interaction prediction, **different** experiments are often aggregated. However, ignoring the potential **discrepancy** between experiments can negatively impact predictive performance.



Affinity

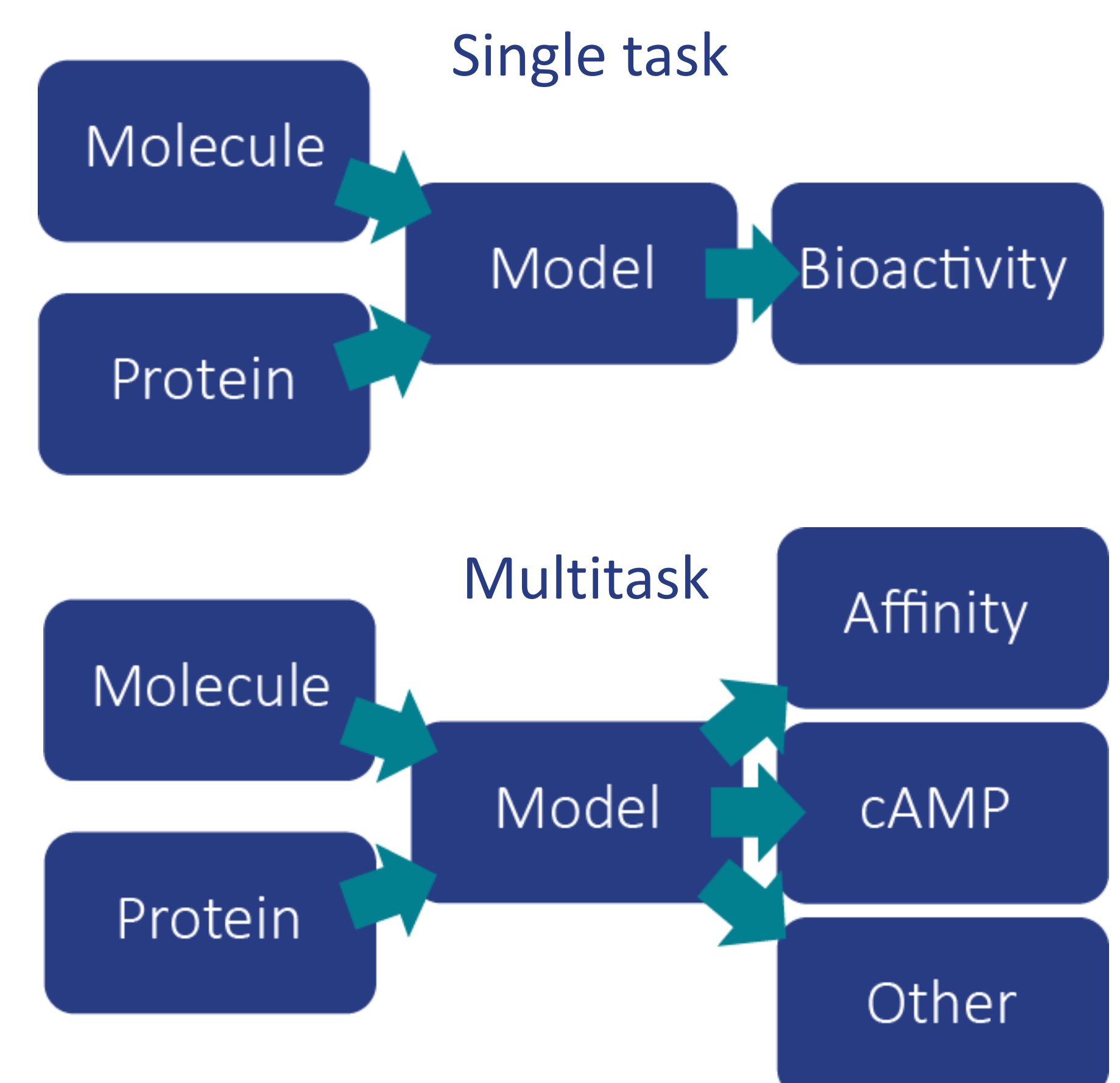


Second messengers



Intracellular changes

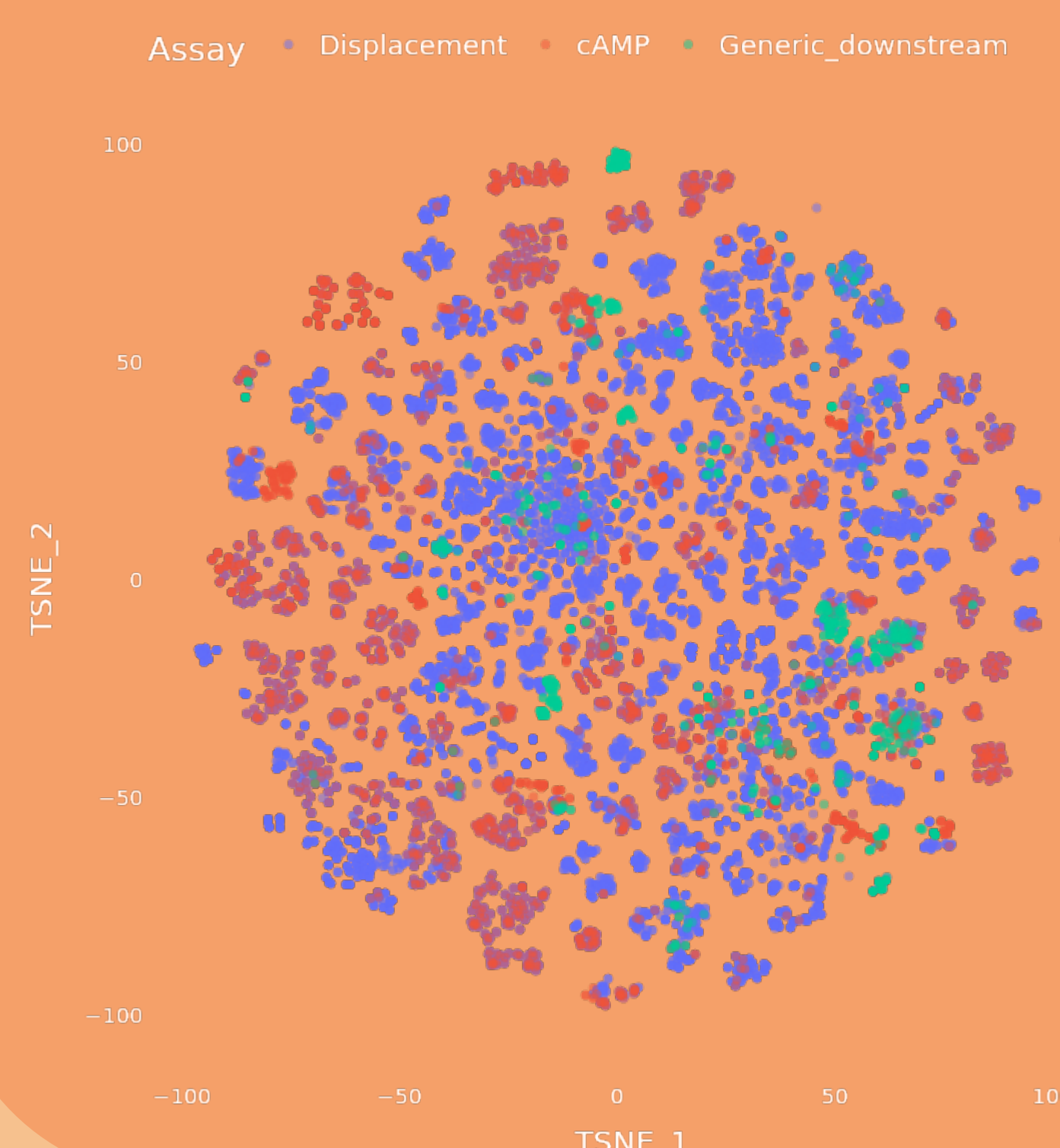
### Building blocks



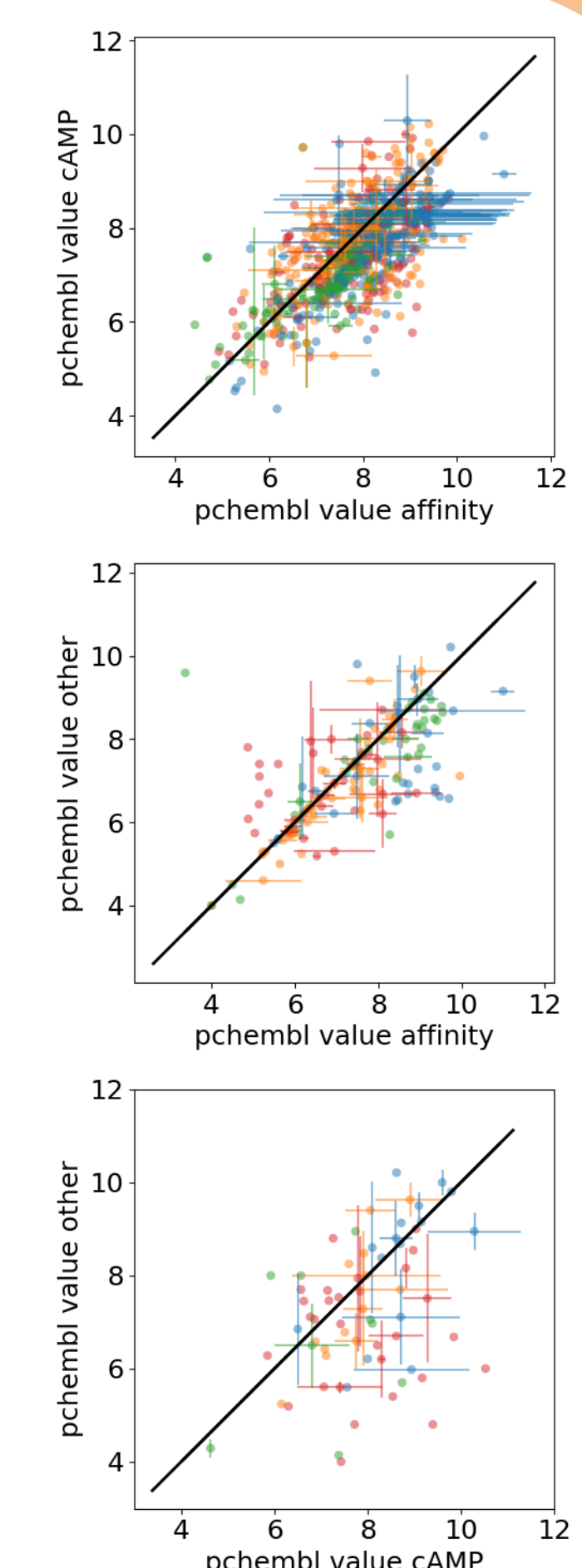
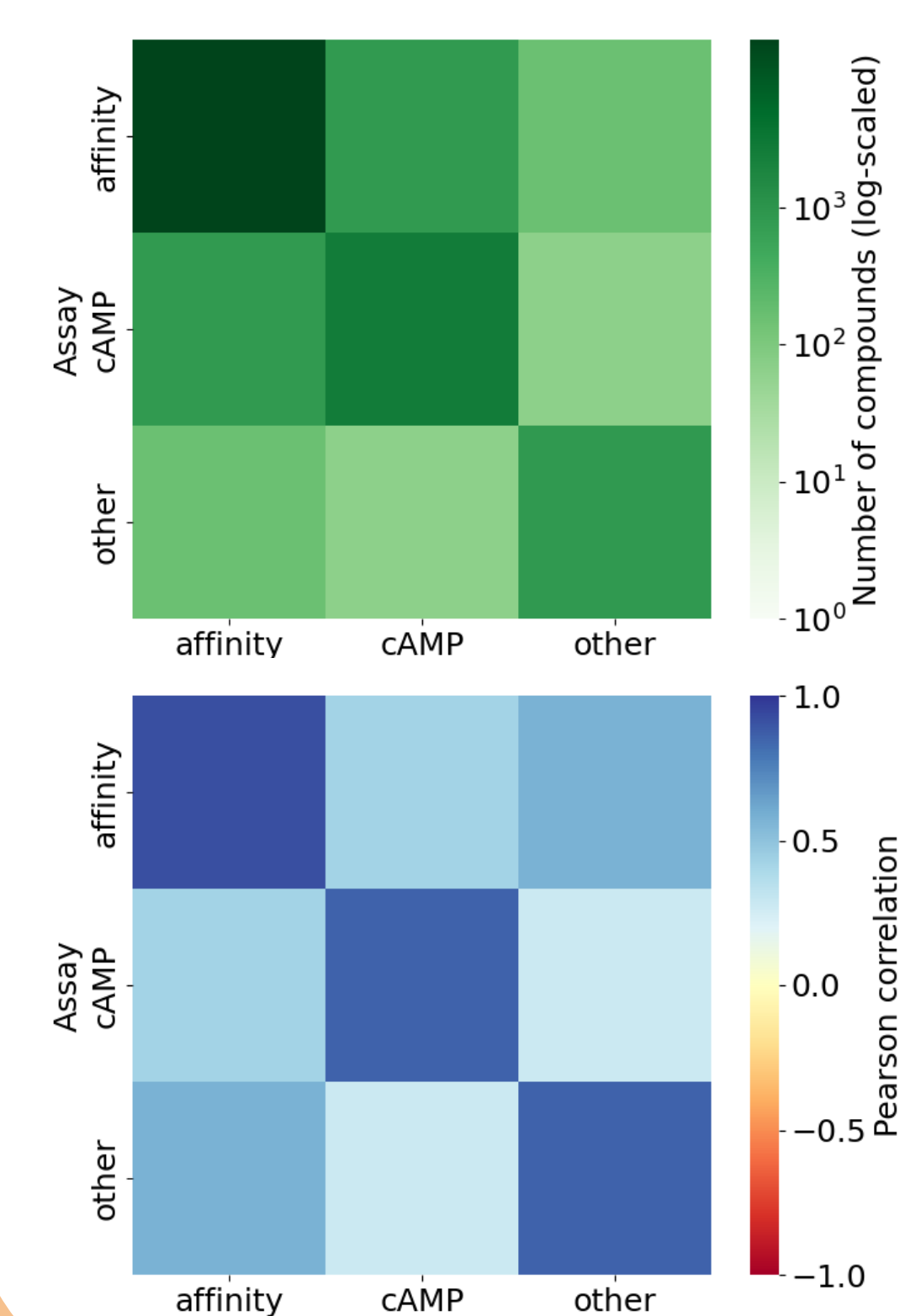
#### Adenosine receptors

- 23k datapoints
- 4 targets
- 9k compounds
- 3 assay types
- 41 % density

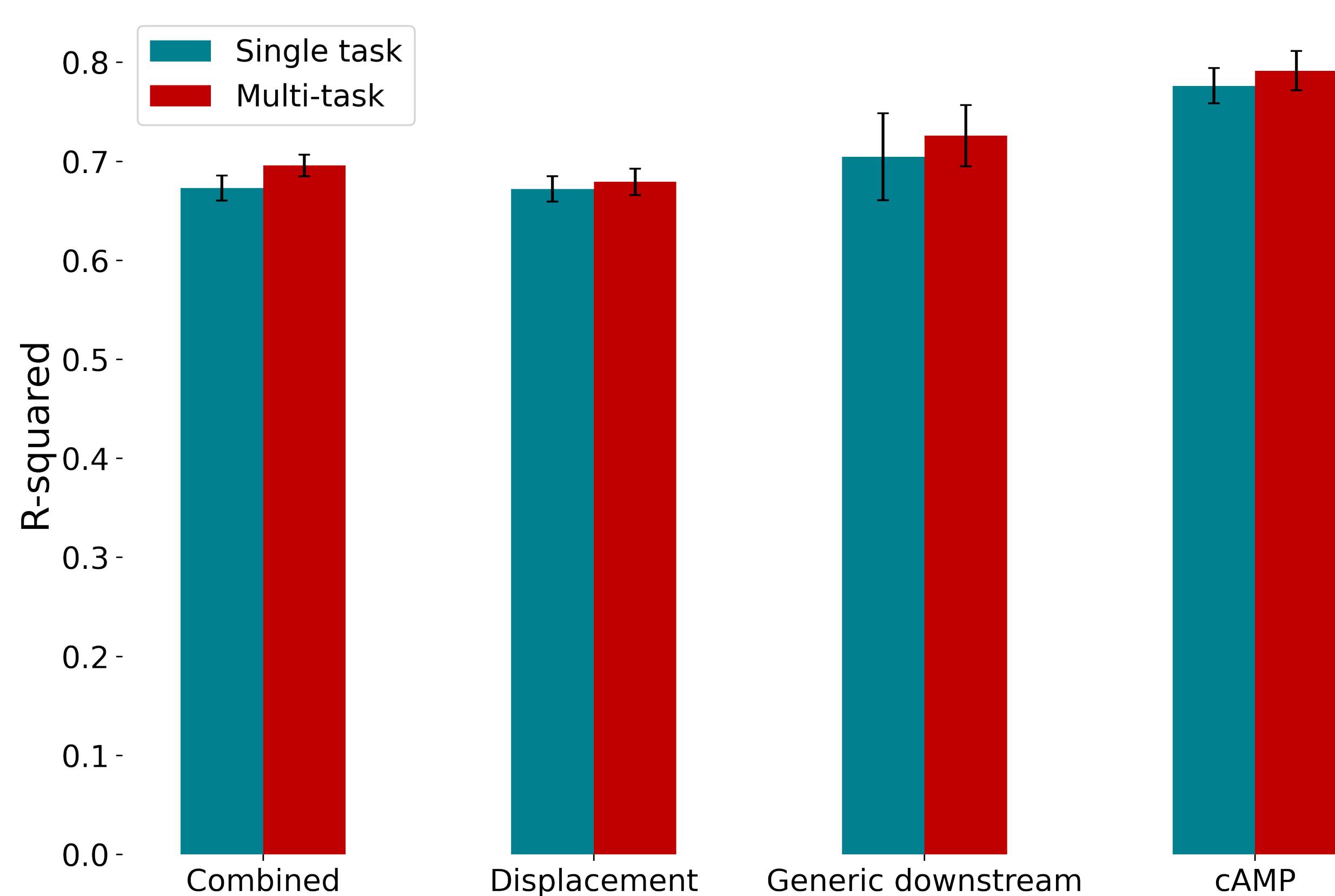
#### Assays in chemical space



#### Assay differences



### Separation increases performance



### Conclusions

Different assay types are **not** fully correlated. Separating them into categories improves the performance of **single task** models. However, there is no significant performance increase for the **multitask** model. These results indicate that this model is not able to take full advantage of the additional data. Further work is needed to investigate, validate and build on these findings.

#### References

Virtual Human Platform for Safety Assessment. <https://vhp4safety.nl/>  
Kalliokoski T. *et al.* 2013, PLoS ONE DOI: 10.1371/journal.pone.0061007  
Pentina, A., Clevert, D. 2022 ChemRxiv. DOI: 10.26434/chemrxiv-2022-d5tzd

