UNIVERSITY OF CAMBRIDGE

# Machine learning: *a priori* or *a posteriori*?

Gareth Conduit

# Machine learning algorithm to

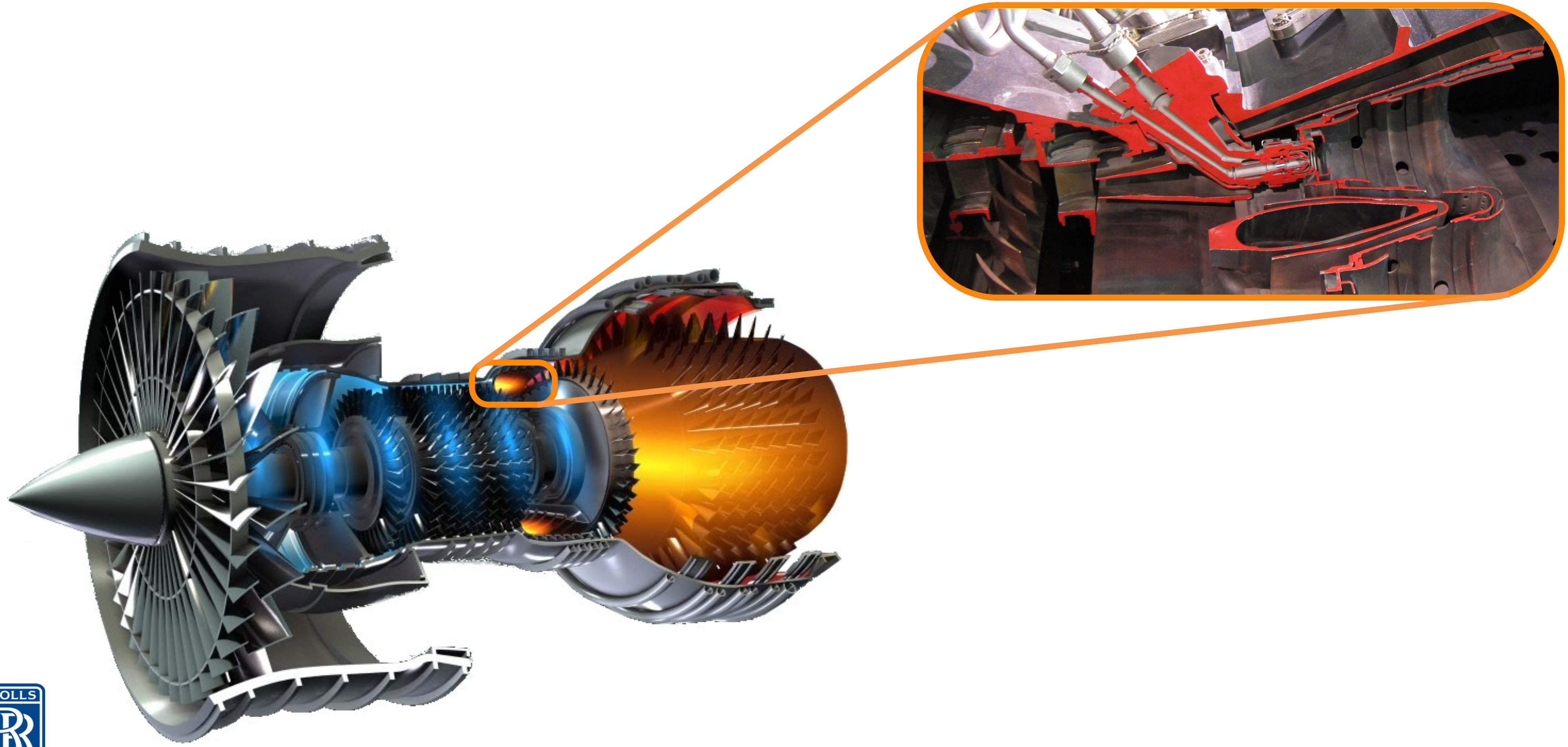
Merge *a priori* computer simulations and physical laws with *a posteriori* experimental data

Exploit *a priori* property-property correlations

Train from sparse datasets

Reduce costly experiments to accelerate discovery

# Combustor in a jet engine

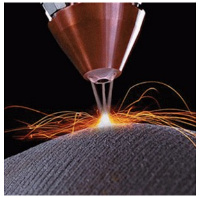


# *A posteriori* black box machine learning for materials design

Properties

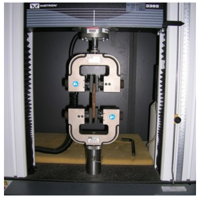
Composition

Defects

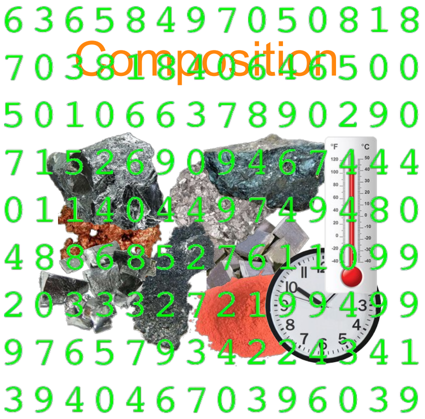


Fatigue

Strength



# Train the *a posteriori* machine learning

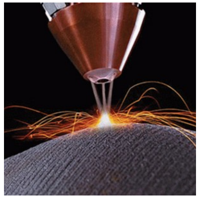



# *A posteriori* machine learning predicts material properties

Composition

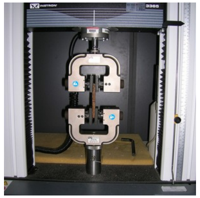
Properties

Defects

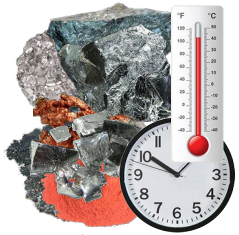


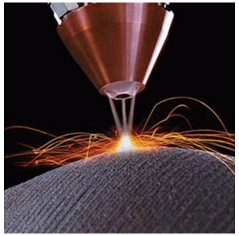
Fatigue

Strength



# Data available to model defect density



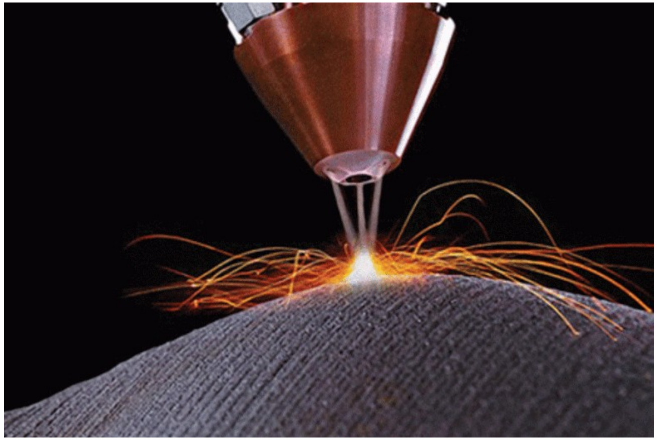


Composition and heat treatment space 30 dimensions

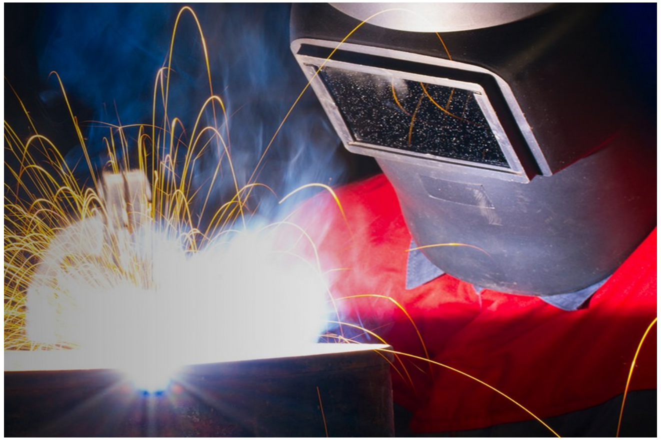
Requires 31 points to fit a hyperplane

Just 10 data entries available to model defect density

# Ability for printing and welding are strongly correlated

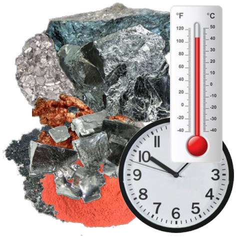


Laser



Electricity

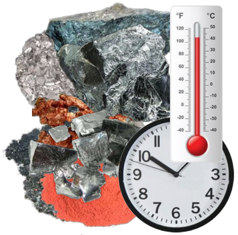
# First predict weldability



1000 entries

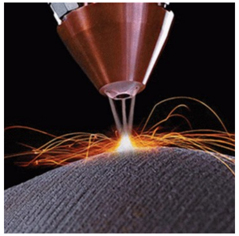
Use 1000 weldability entries to understand complex composition → weldability model

# Use *a posteriori* weldability to *a priori* predict defects formed



1000 entries

10 entries

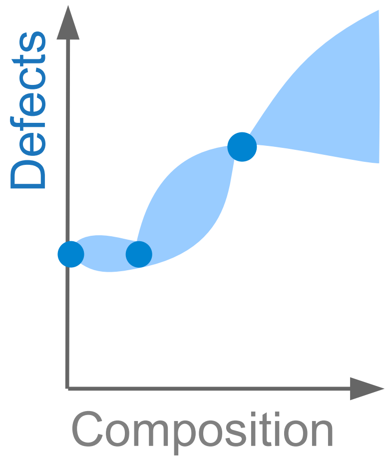


Use 1000 weldability entries to understand complex composition → weldability model

10 defects entries capture the simple weldability → defect relationship

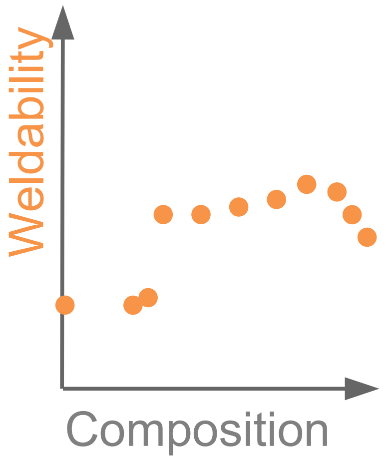
Two interpolations give composition → defects extrapolation

# Insufficient data for number of defects formed

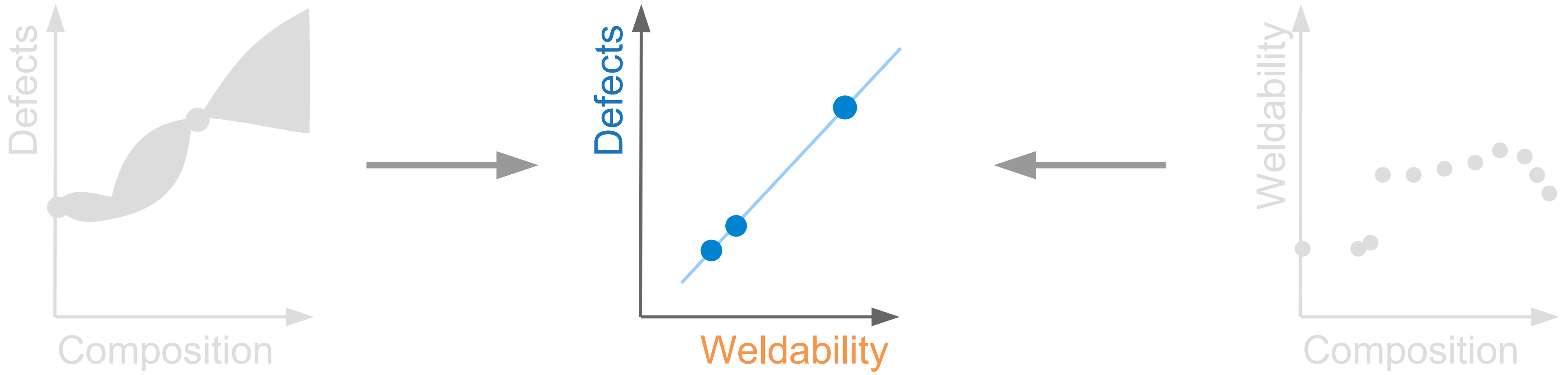


# Welding is analogous to direct laser deposition

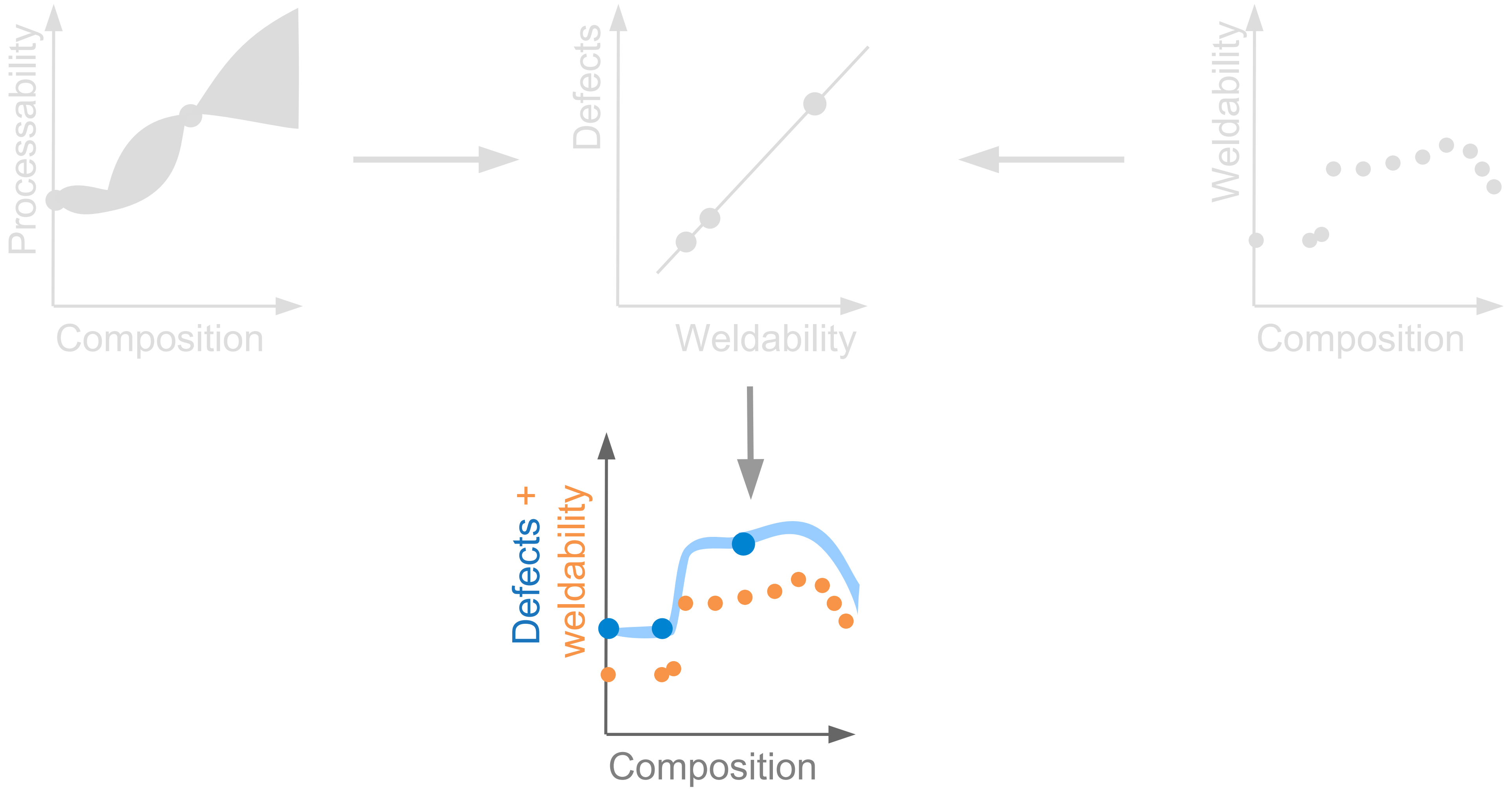


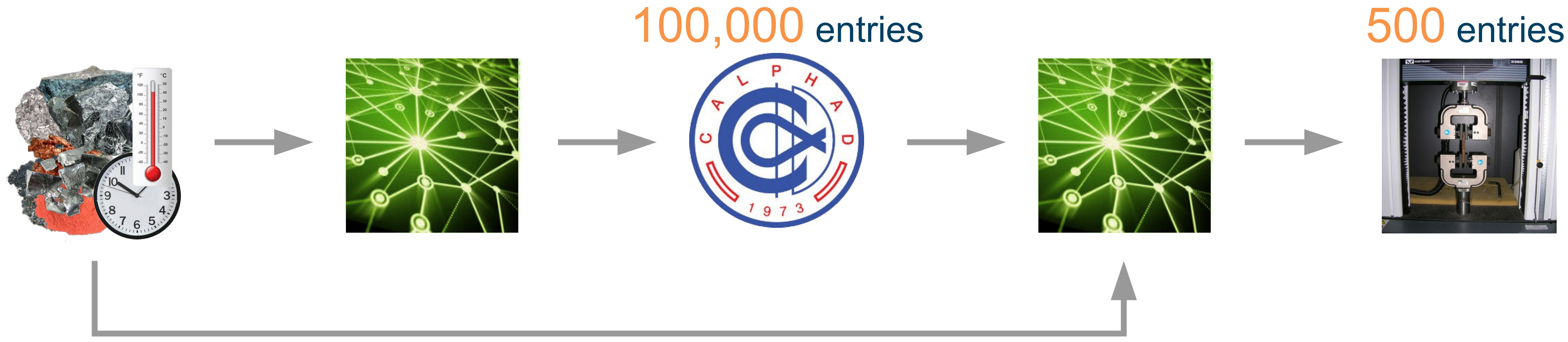
# Straightforward defects-welding relationship



# Merge properties with the machine learning



# Use *a priori* CALPHAD to *a priori* predict strength




Use 100,000 CALPHAD results to model complex composition → phase behavior

500 strength entries capture the phase behavior → strength relationship

Two interpolations aid the composition → strength extrapolation

# Target properties

| | | |
|---|---|---|
| Elemental cost | < | $25\ \$\mathrm{kg}^{-1}$ |
| Density | < | 8500 $\mathrm{kgm}^{-3}$ |
| Defects | < | 0.15% defects |
| Oxidation resistance | < | 0.3 $\mathrm{mgcm}^{-2}$ |
| γ content | > | 75 wt% |
| Phase stability | > | 99 wt% |
| γ’ solvus | > | 1000˚C |
| Thermal resistance | > | 0.04 $\mathrm{K\Omega^{-1}m^{-3}}$ |
| Yield stress at 900˚C | > | 200 MPa |
| Tensile strength at 900˚C | > | 300 MPa |
| Tensile elongation at 700˚C | > | 8% |
| 1000hr stress rupture at 800˚C | > | 100 MPa |
| Fatigue life at 500 MPa, 700˚C | > | $10^5$ cycles |

# Composition and processing variables

Cr 19%

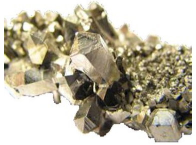
Co 4%

Mo 4.9%

W 1.2%

Zr 0.05%

Nb 3%



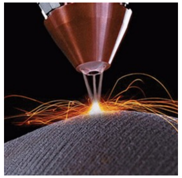
Al 2.9%

C 0.04%

B 0.01%

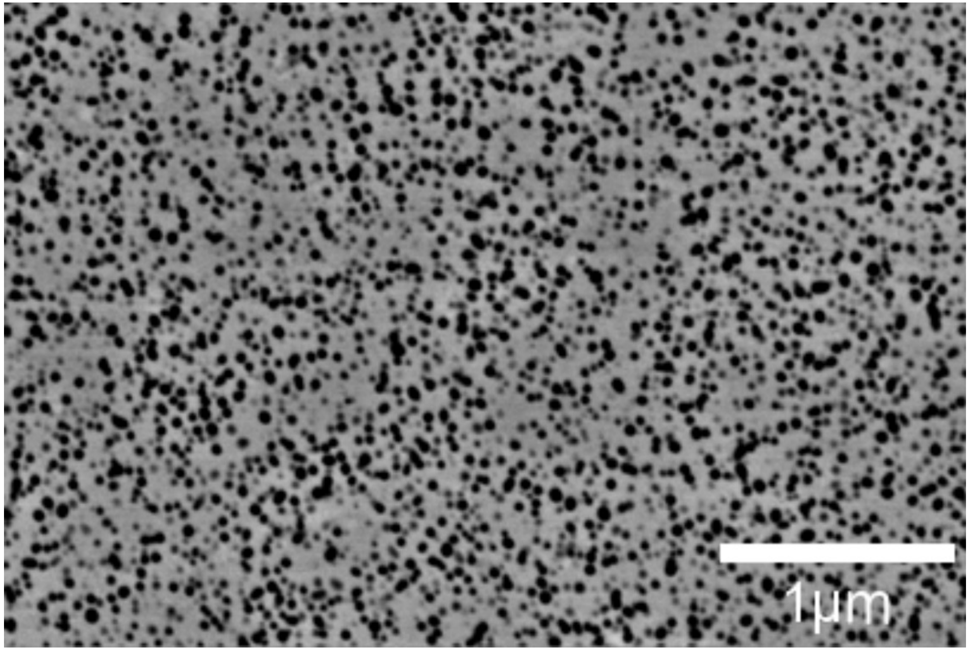
Ni

Expose 0.8



$T_{HT}$ 1300ºC

# Microstructure




Probabilistic neural network identification of an alloy for direct laser deposition
Materials & Design 168, 107644 (2019)

# Target γ content

Elemental cost < 25 $\$kg^{-1}$

Density < 8500 $kgm^{-3}$

Defects < 0.15% defects

Oxidation resistance < 0.3 $mgcm^{-2}$

γ content > 75 wt%

Phase stability > 99 wt%

γ' solvus > 1000˚C

Thermal resistance > 0.04 $K\Omega^{-1}m^{-3}$

Yield stress at 900˚C > 200 MPa

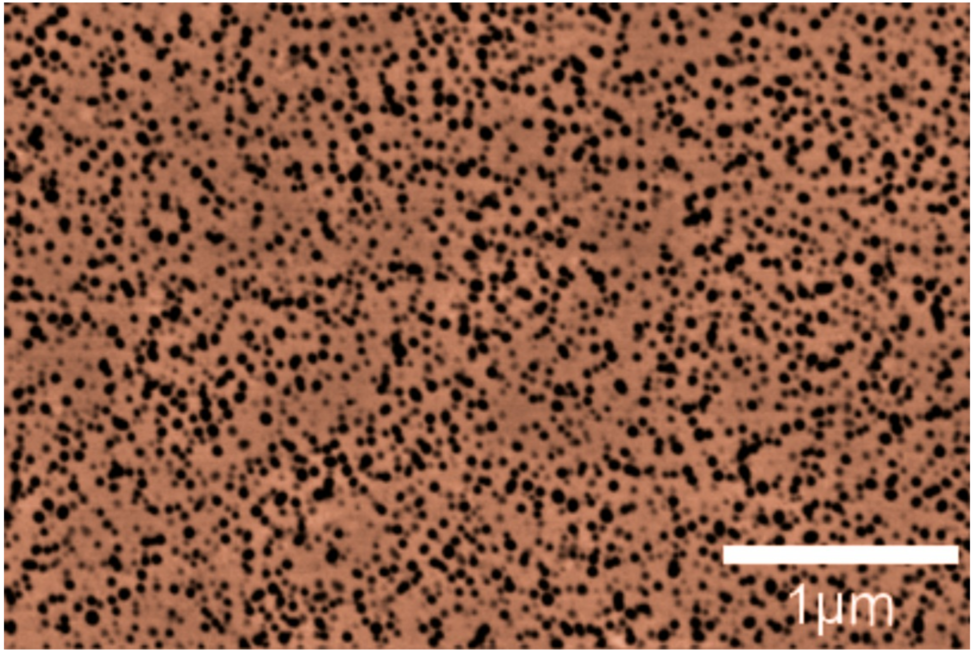
Tensile strength at 900˚C > 300 MPa

Tensile elongation at 700˚C > 8%

1000hr stress rupture at 800˚C > 100 MPa

Fatigue life at 500 MPa, 700˚C > $10^5$ cycles

# Microstructure




Probabilistic neural network identification of an alloy for direct laser deposition
Materials & Design 168, 107644 (2019)

# Target phase stability

Elemental cost $< 25\ \$\mathrm{kg}^{-1}$

Density $< 8500\ \mathrm{kgm}^{-3}$

Defects $<$ 0.15% defects

Oxidation resistance $< 0.3\ \mathrm{mgcm}^{-2}$

γ content $>$ 75 wt%

**Phase stability $>$ 99 wt%**

γ’ solvus $>$ 1000˚C

Thermal resistance $> 0.04\ \mathrm{K\Omega^{-1}m^{-3}}$

Yield stress at 900˚C $>$ 200 MPa

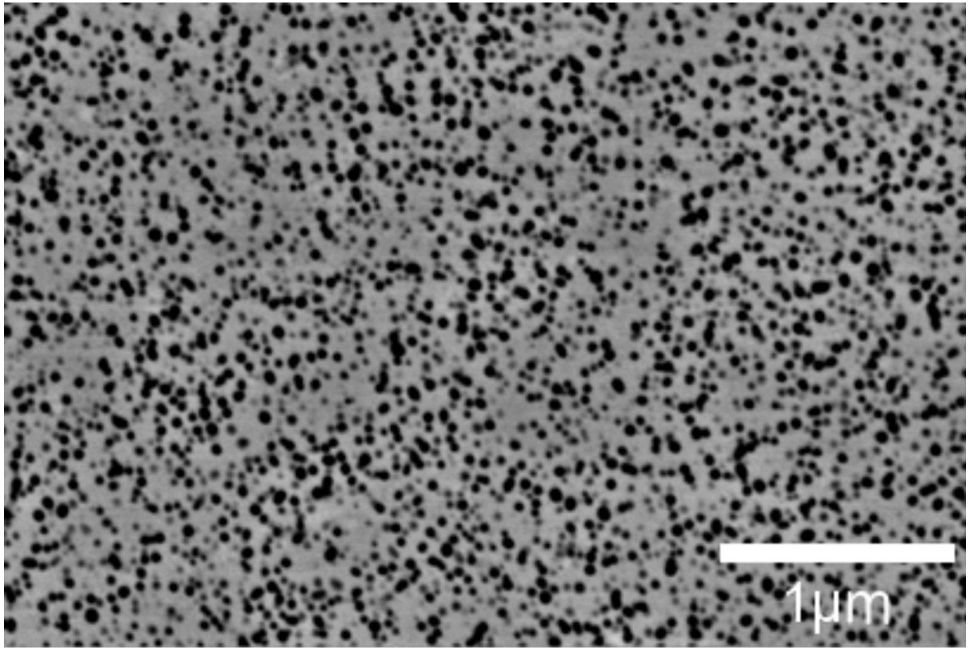
Tensile strength at 900˚C $>$ 300 MPa

Tensile elongation at 700˚C $>$ 8%

1000hr stress rupture at 800˚C $>$ 100 MPa

Fatigue life at 500 MPa, 700˚C $> 10^5$ cycles

# Deleterious phases formed




Probabilistic neural network identification of an alloy for direct laser deposition
Materials & Design 168, 107644 (2019)

# Target defect density

Elemental cost $< 25\ \$\mathrm{kg}^{-1}$

Density $< 8500\ \mathrm{kgm}^{-3}$

Defects $<$ 0.15% defects

Oxidation resistance $< 0.3\ \mathrm{mgcm}^{-2}$

γ content $>$ 75 wt%

Phase stability $>$ 99 wt%

γ’ solvus $>$ 1000˚C

Thermal resistance $> 0.04\ \mathrm{K\Omega^{-1}m^{-3}}$

Yield stress at 900˚C $>$ 200 MPa

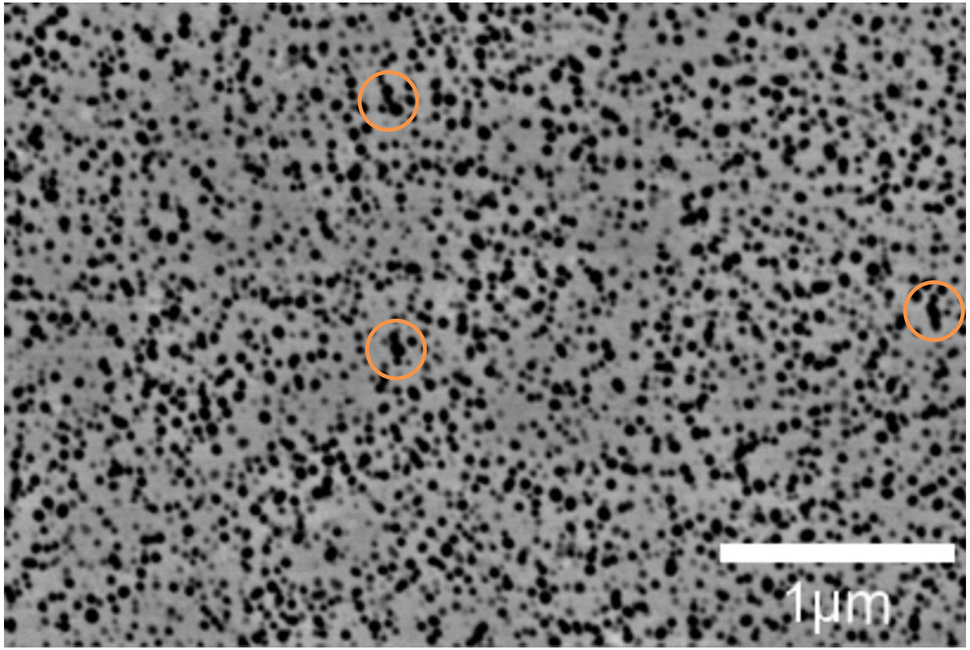
Tensile strength at 900˚C $>$ 300 MPa

Tensile elongation at 700˚C $>$ 8%

1000hr stress rupture at 800˚C $>$ 100 MPa

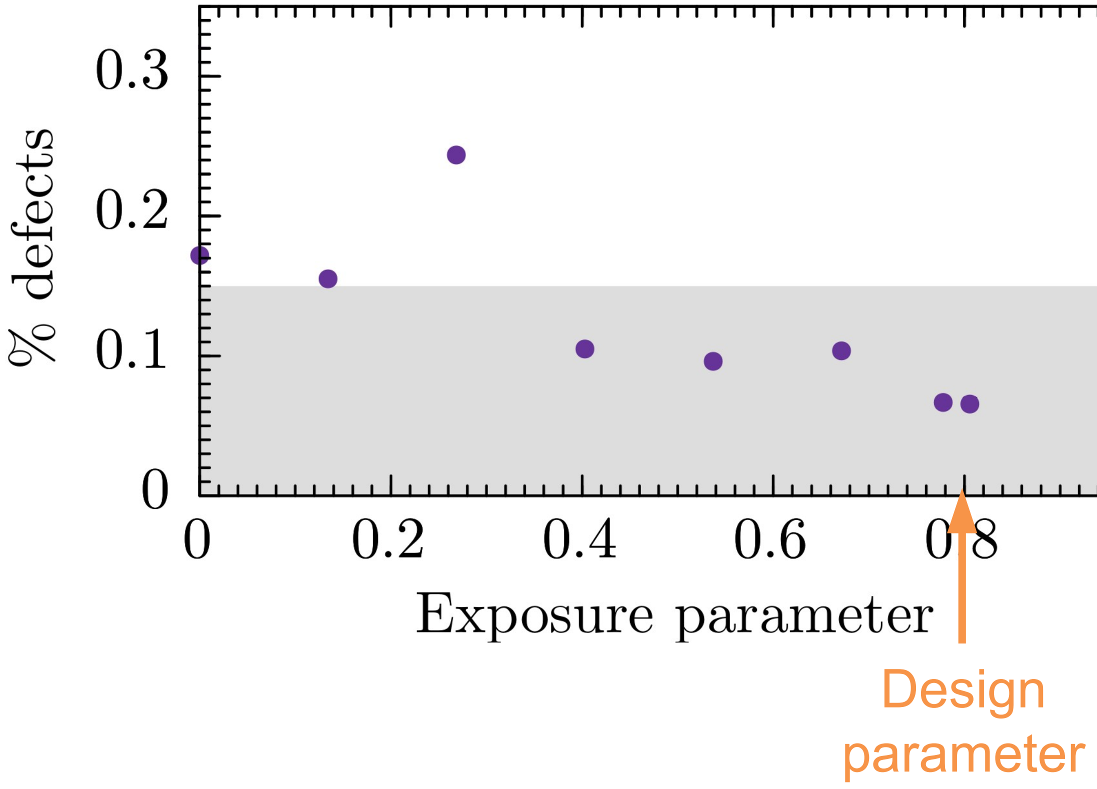
Fatigue life at 500 MPa, 700˚C $> 10^5$ cycles

# Defect detection




Probabilistic neural network identification of an alloy for direct laser deposition
Materials & Design 168, 107644 (2019)

# Testing the defect density




Probabilistic neural network identification of an alloy for direct laser deposition
Materials & Design 168, 107644 (2019)

DOMINO

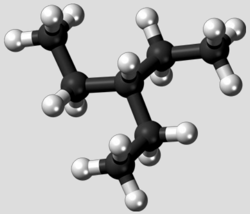
Ordnance Survey

The University Of Sheffield. AMRC Advanced Manufacturing Research Centre

BOEING

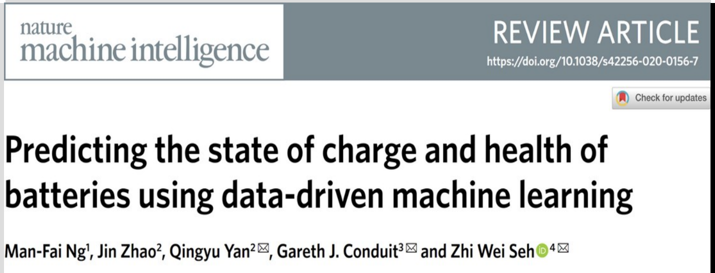
GKN

Agency for Science, Technology and Research SINGAPORE



bp

Lubricants
for electric cars



nature machine intelligence

REVIEW ARTICLE

https://doi.org/10.1038/s42256-020-0156-7

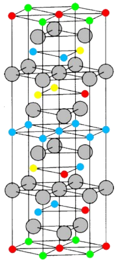
Check for updates

# Predicting the state of charge and health of batteries using data-driven machine learning

Man-Fai Ng[1], Jin Zhao[2], Qingyu Yan[2]✉, Gareth J. Conduit[3]✉ and Zhi Wei Seh[4]✉

Agency for Science, Technology and Research SINGAPORE

SAMSUNG

NASA

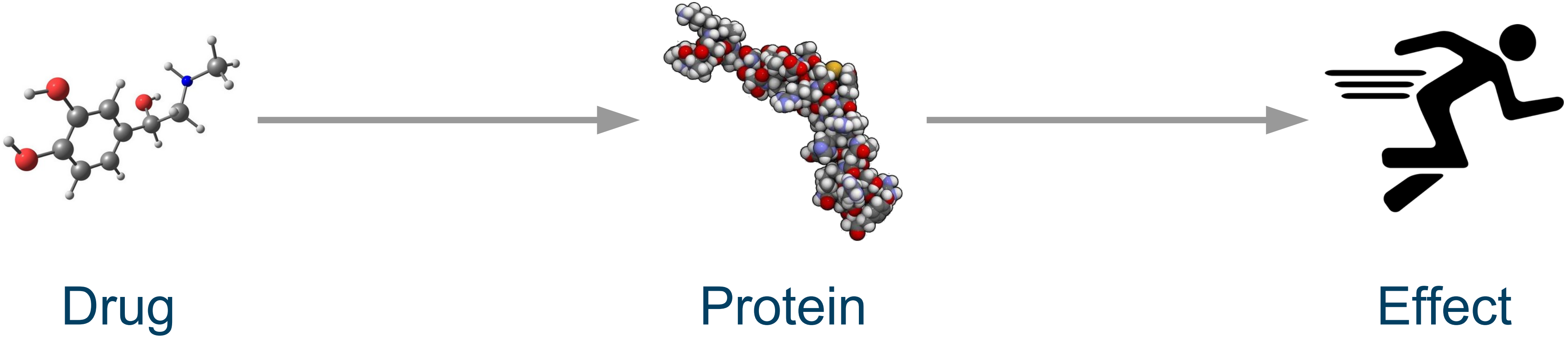
Heat exchanger
& shape memory
alloy applications

# Open Source Malaria contest

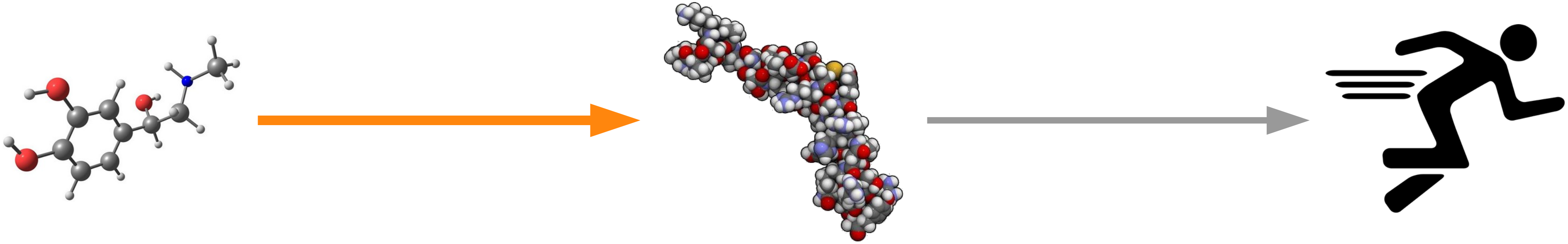
OPEN SOURCE MALARIA

Looking for New Medicines

# Action of a drug

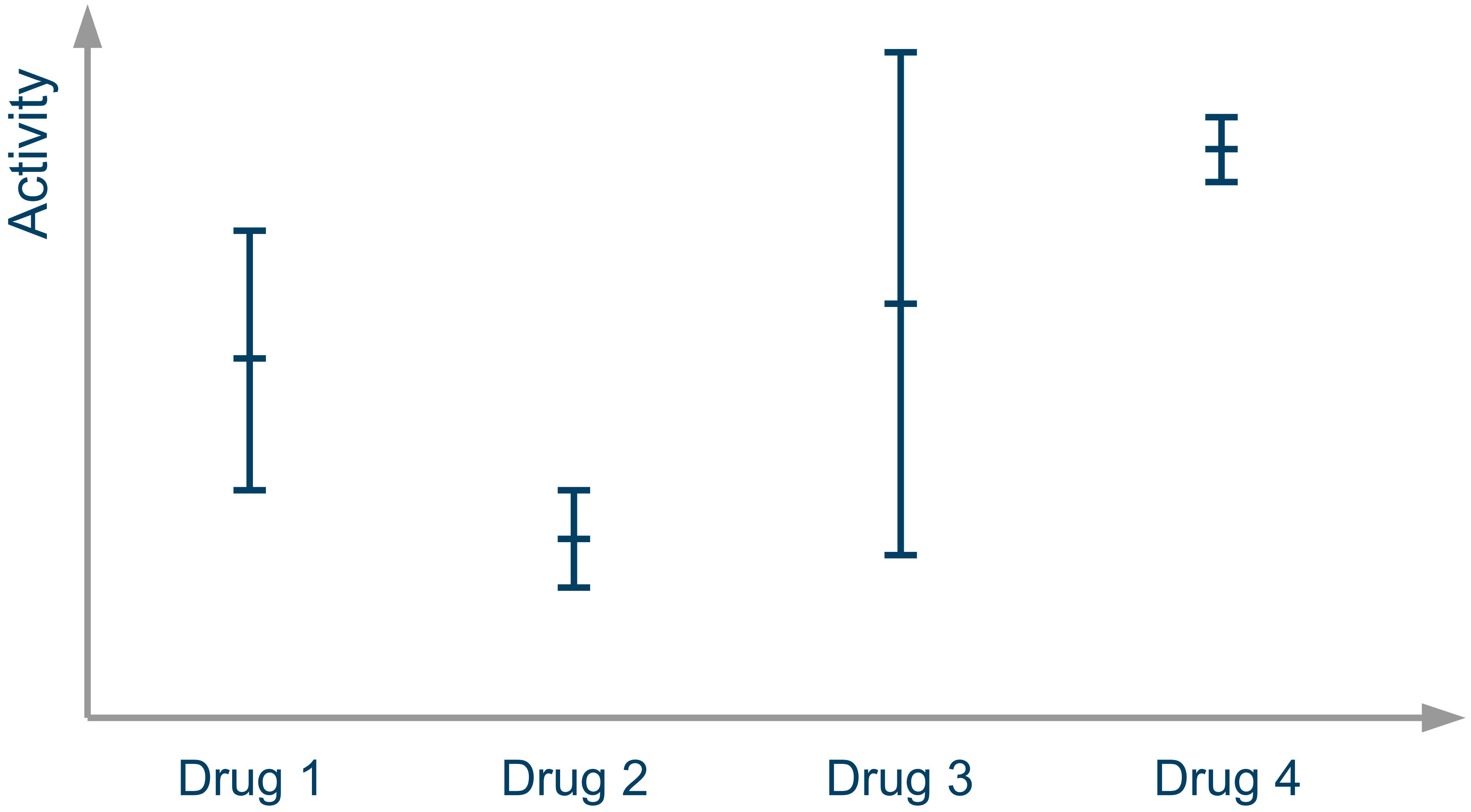


# Action of a drug

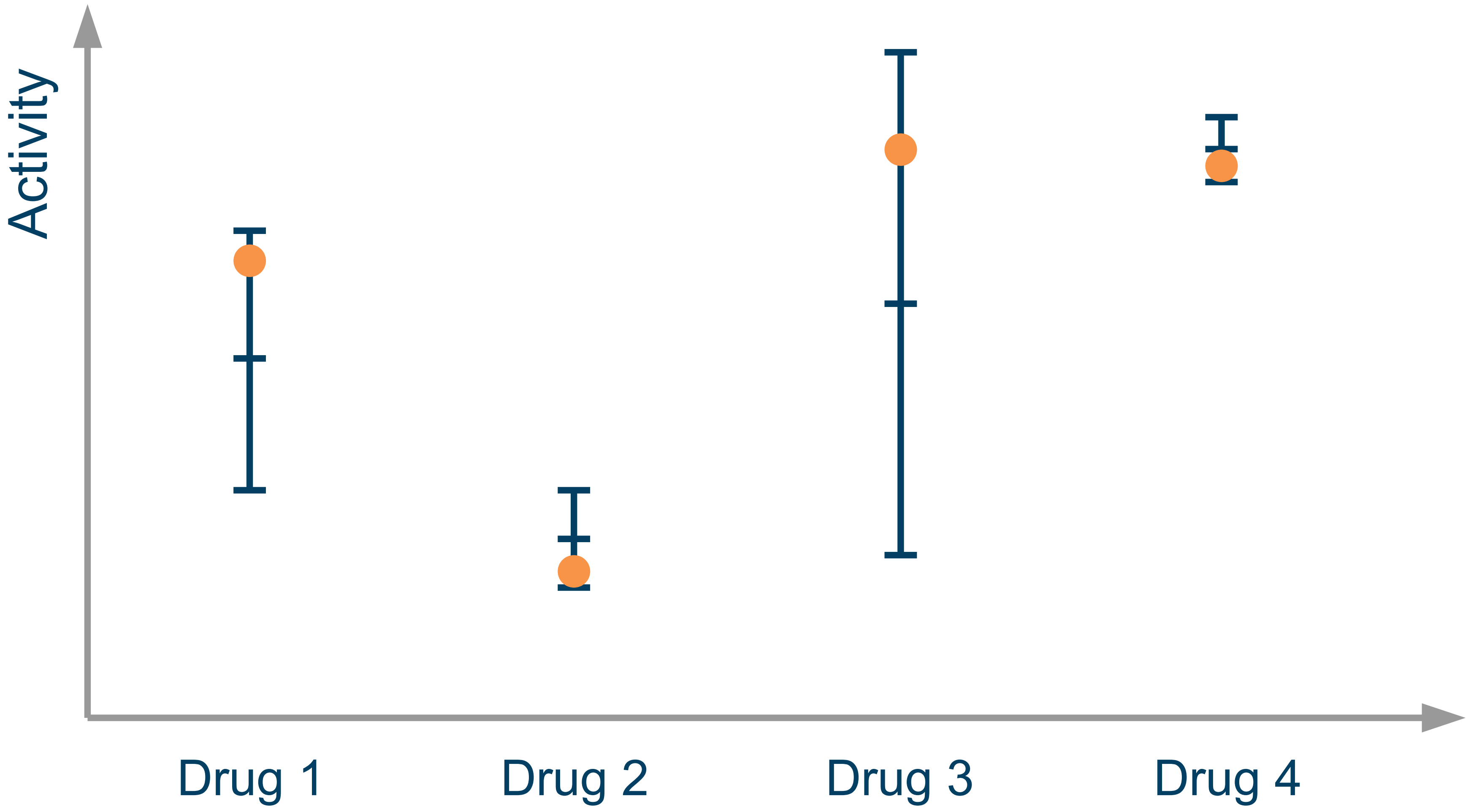


Drug | Protein | Effect

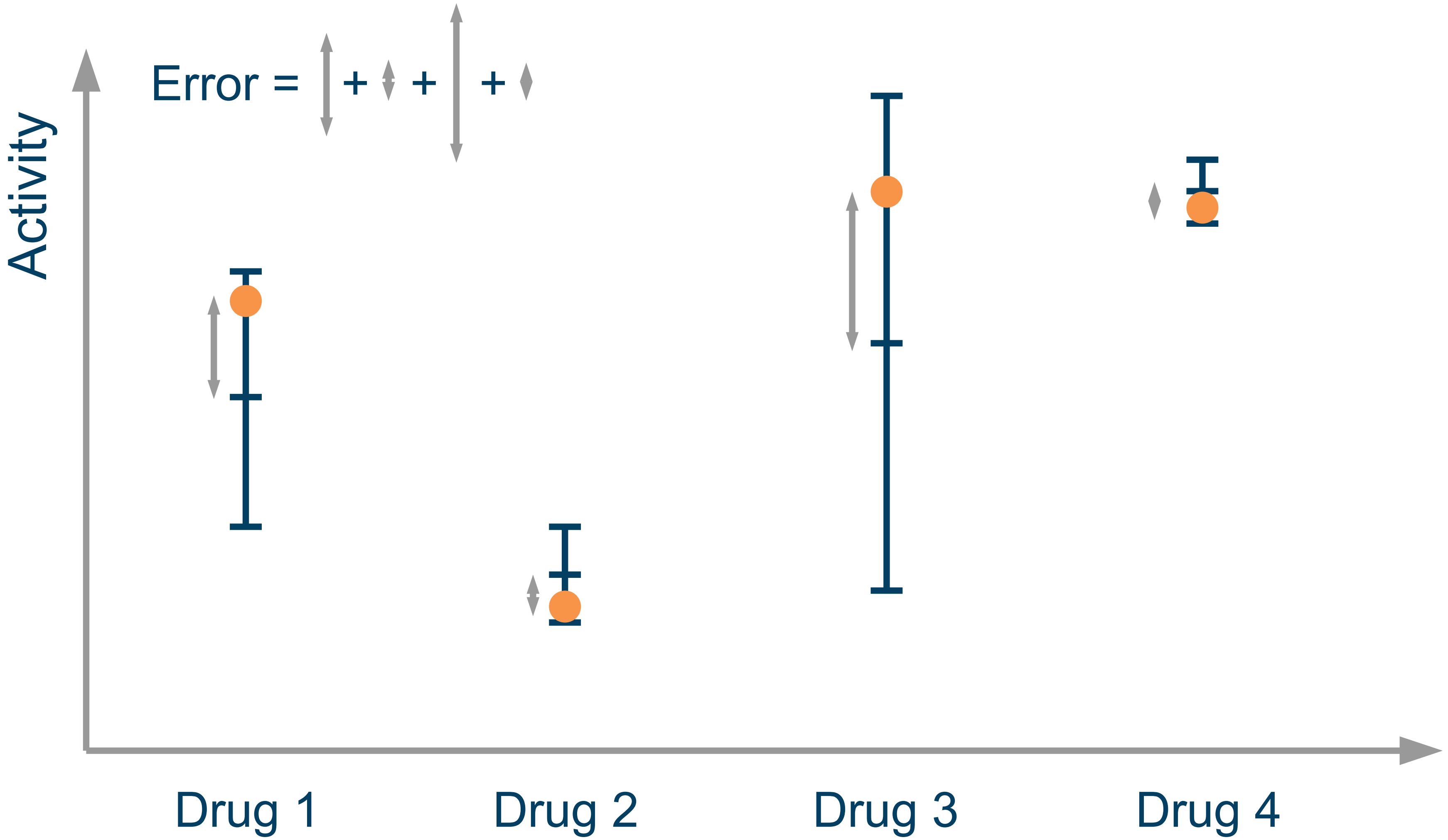
# Predictions have an uncertainty



# Validation data typically within one standard deviation

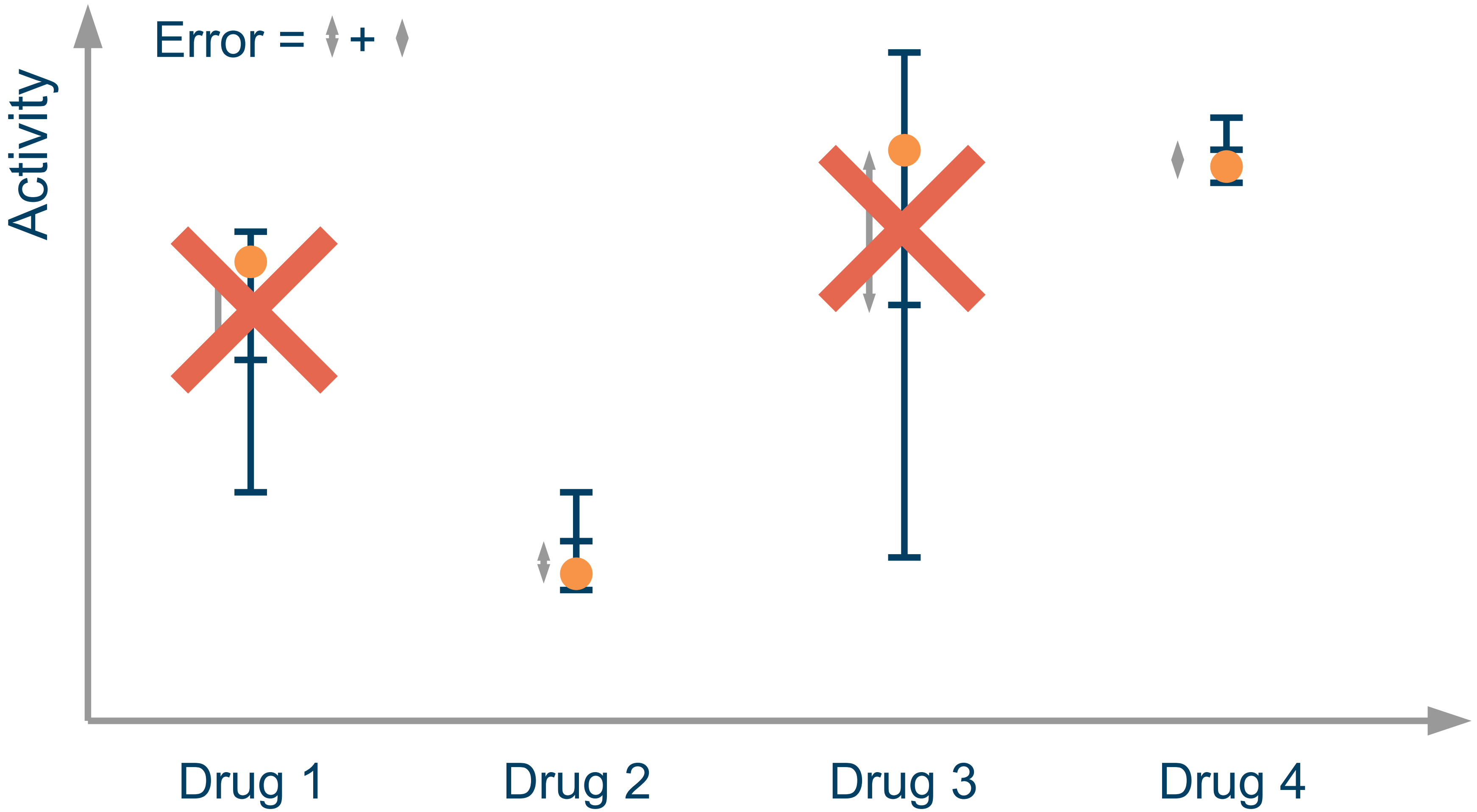


# $R^2$ metric calculated with difference from predicted value

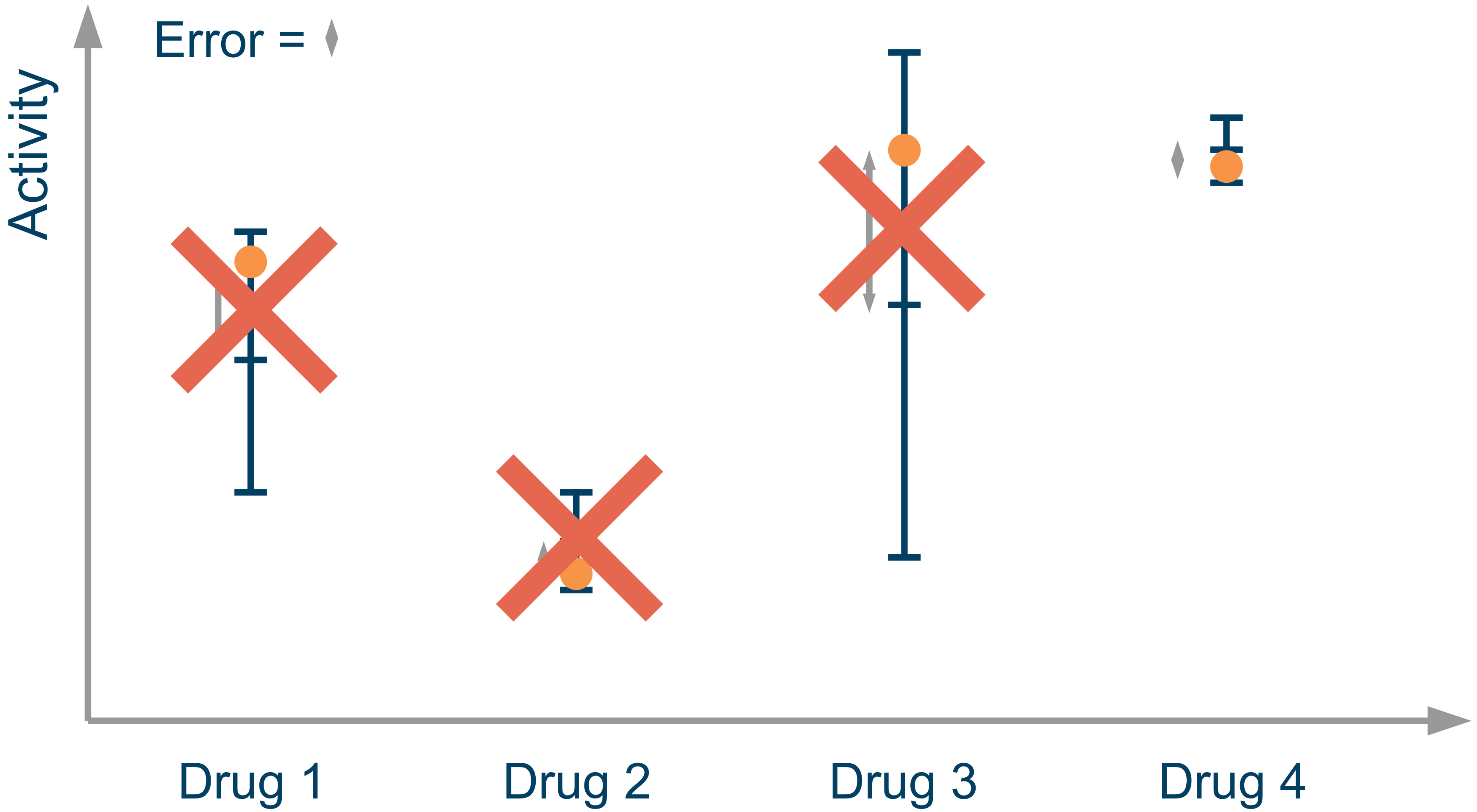


# Impute 75% of data with smallest uncertainty

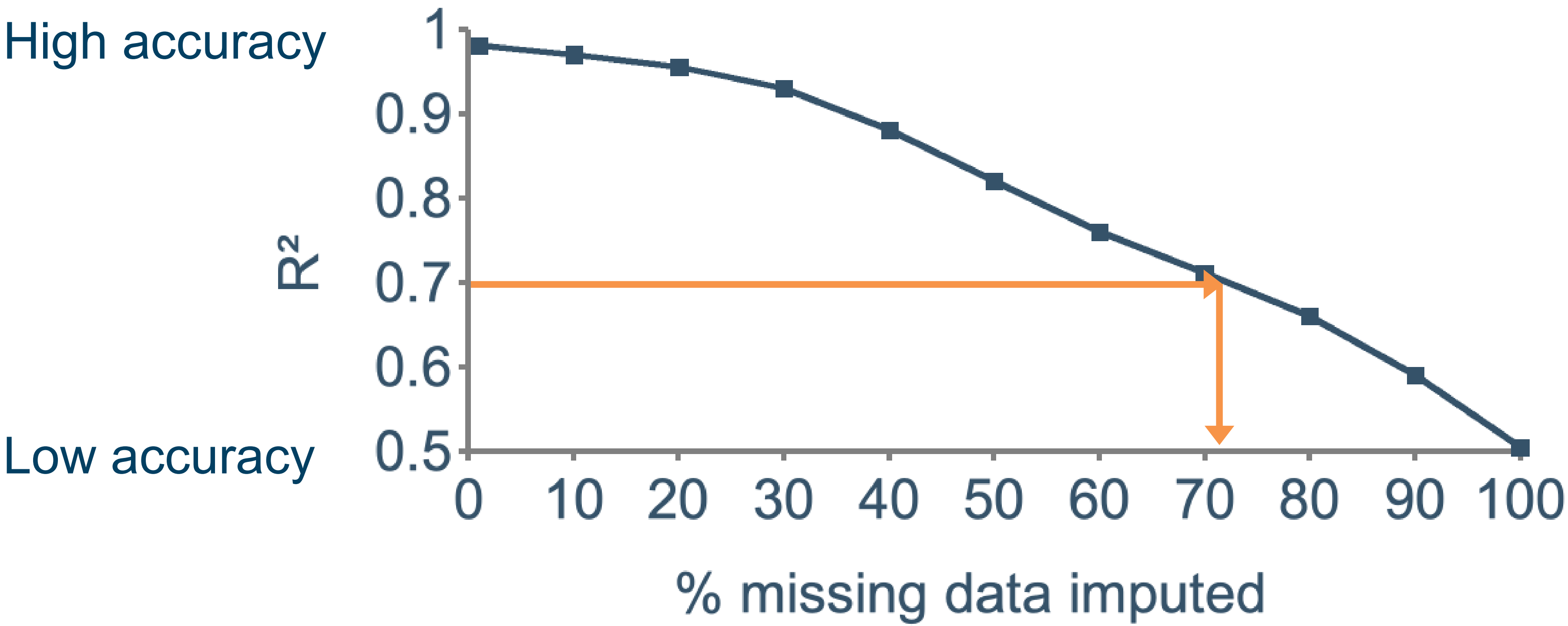
# Impute 50% of data with smallest uncertainty



# Impute 25% of data with smallest uncertainty

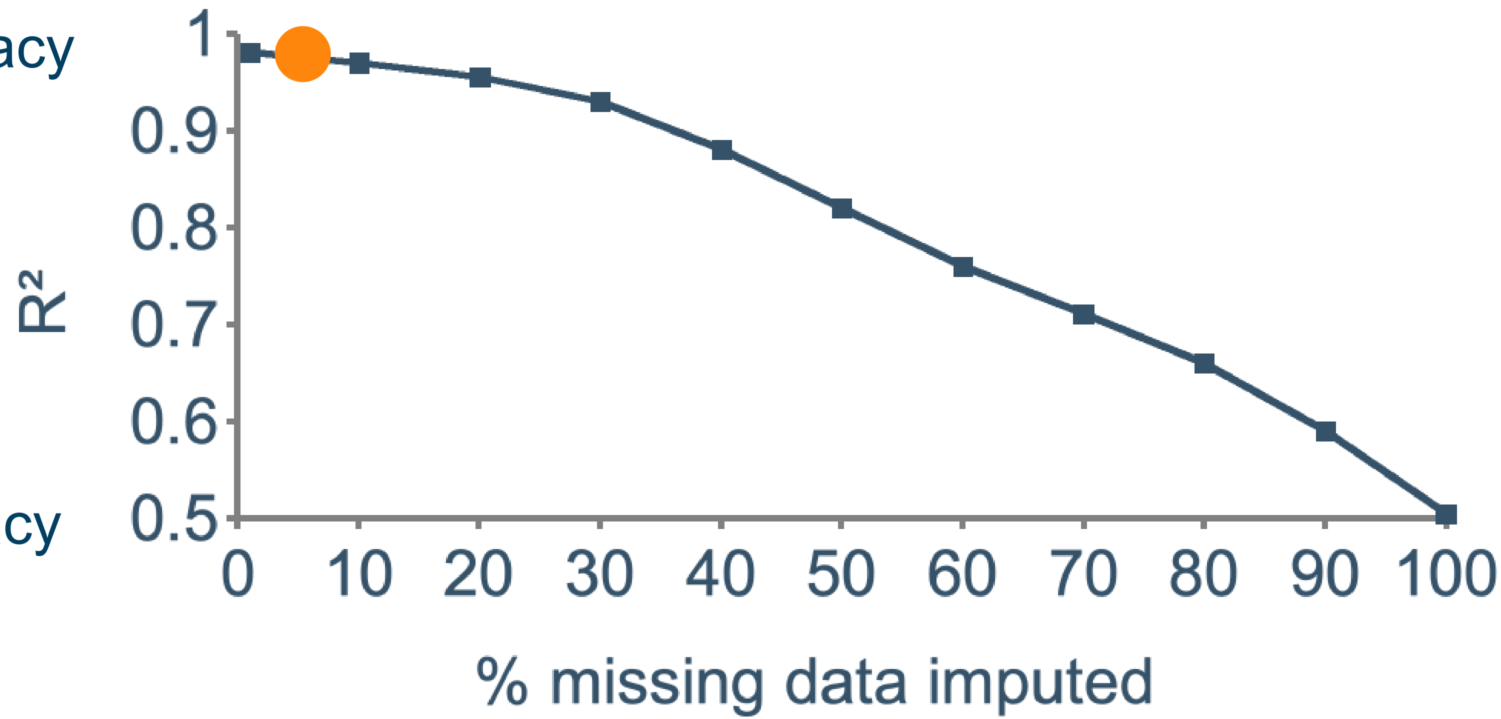


# Improved performance by exploiting uncertainty



# Focus on compounds with low uncertainty

High accuracy

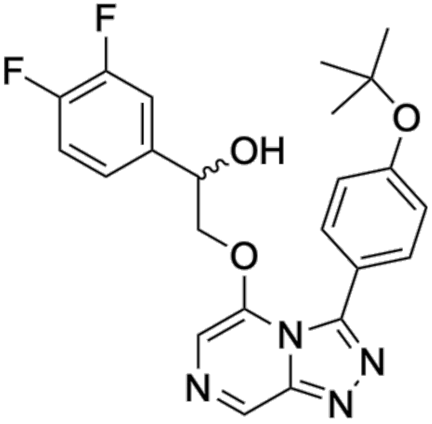



Low accuracy

# Different drugs can treat the same ailment




# Open Source Malaria experimental validation



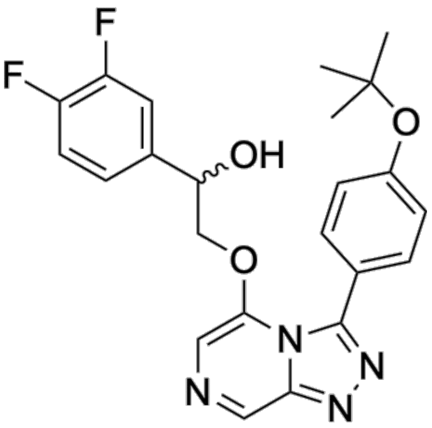

Optibrium & Intellegens | Davy Guan | Exscientia | Molomics

0.647 μM

Journal of Medicinal Chemistry 64, 16450 (2021)

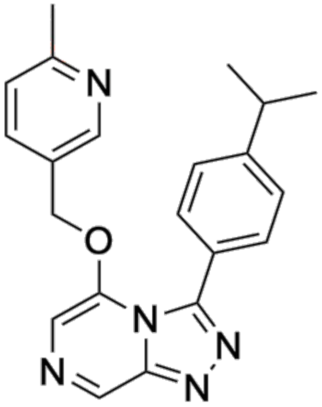
OPEN SOURCE MALARIA
Looking for New Medicines

# Open Source Malaria other compounds



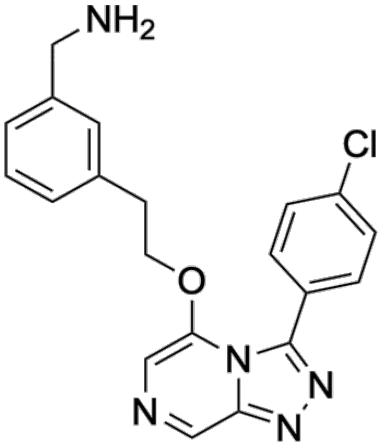

Optibrium & Intellegens

0.647 μM



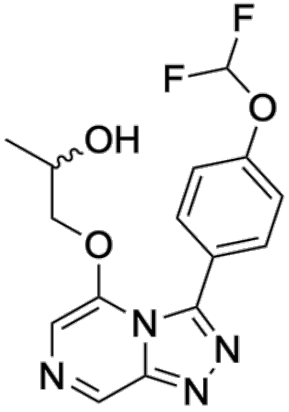

Davy Guan

>25 μM




Exscientia

10.9 μM




Molomics

>25 μM

Journal of Medicinal Chemistry 64, 16450 (2021)

OPEN SOURCE MALARIA
Looking for New Medicines

# Commercialization

intellegens

Alchemite Analytics™ platform for materials and chemicals with Intellegens released in September 2020

optibrium

Machine learning tool embedded into Cerella™ released in October 2020

Ansys / GRANTA

Machine learning integrated into Granta Material Intelligence™ released in January 2022

# Summary

Merge simulation with experimental data and exploit property-property relationships to circumvent missing data, designed an experimentally verified alloy for 3d printing

Exploited uncertainty to predict drug most probable drug

Generic approach applied to materials, batteries, pharmaceuticals, and beyond

Taken to market through startup Intellegens as Alchemite Analytics™ and with partners Optibrium and Ansys

intellegens optibrium Ansys / GRANTA