

Khan I^{1,2}, Crott R², Martina R², Begum R², Khan Y²

¹University of Warwick, Coventry, United Kingdom; ²Regulatory Scientific and Health Solutions (R-S-S), Shirley, Solihull, United Kingdom.

INTRODUCTION

We investigated the application of machine learning (ML) methods through symbolic regression (SR) to generate improved model fits that are not humanly possible within the time constraints of submission. We show utilities from ML algorithms can impact significantly the incremental cost-effectiveness ratio (ICER) and the quality adjusted life year (QALY) estimates.

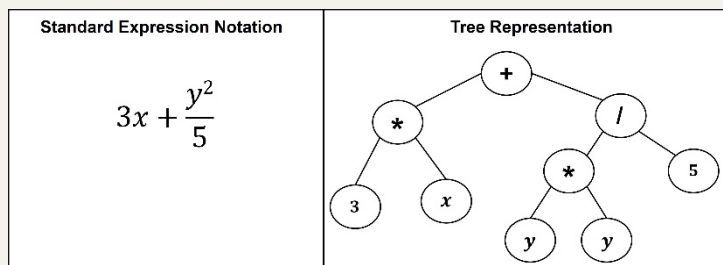
METHODS

We use data from a previously reported randomized (Erlotinib vs Best Supportive Care (BSC) trial (TOPICAL)¹ in stage IIb-IV non-small cell lung cancer (NSCLC) patients (N=670). Health states were defined as progression free (PF), progressive disease (PD) and Death. EQ-5D-3L utilities were collected at baseline and monthly until disease progression.

Symbolic Regression

Symbolic regression (SR) attempts to fit an equation through a set of observed data points³. In general, equations (model form) need to be identified first, prior to fit. Through SR, we can find the equation(s) that fit the sample data through an optimization criteria (e.g. R², Aikakes information criterion (AIC)). Genetic Programming (GP) is an implementation of evolutionary programming (**Figure 1**), where the problem-solving domain is modelled on a computer and the algorithm attempts to find a solution by the process of simulated evolution, employing the biological theory of genetics and the Darwinian principle of survival of the fittest³. In this sense, GP is an ideal vehicle to implement SR. Data Modeler® using a Mathematica® platform utilizing all mathematical operators except step, hyperbolic and trigonometric functions were used.

Figure 1: Description of GP



Source: Hussein. S³

RESULTS

Trial Demographics: These have been published elsewhere^{1,2}; In general, N=350 vs 320 Erlotinib vs BSC; median age 77 yrs, 61% male; 35% stage IIb; ECOG 0-1,2,3: 16%, 41%, 43% respectively.

Table 1: Utilities: Summary Statistics

Health State	Mean EQ-5D (SE)		
	Observed	Linear Model	SR#
Alive and PF (n=33)	0.745 (0.019)	0.793 (0.027)	0.744 (0.021)
PD (n=637)	0.612 (0.0061)	0.660 (0.011)	0.611 (0.017)
AIC		318	272 to 290
R ²		4.2%	6.5% - 9.2%
Incremental QALY		0.035 (0.0163)*	0.029 (0.033)

#Mildly complex to Most complex; models included covariates: ECOG, smoking status, age, gender; *published in Khan et al, 2015¹

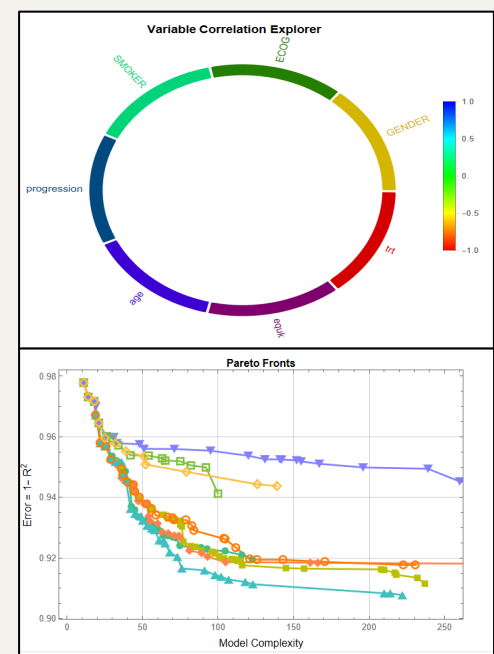
SYMBOLIC REGRESSION MODELS

More Complex SR Model (Level 222)

222	0.908	$\frac{1}{-e^{eCOG^{1/3} + 2eCOG + Smoker \cdot 0.48 + 1.74} - \frac{3428.62}{1} - \frac{3300.82}{age} - 0.33eCOG - e^{eCOG + Smoker + (eCOG + trt)^{1/3} - age} - 0.28Gender - \frac{152.96eCOG}{age + e^{eCOG}ageGender} - 0.15Progression}$

Less Complex SR Model (Level 42)

42	0.936	$0.95 - \frac{28.84}{age} - (6.37 \times 10^{-2})eCOGProgression + \frac{5.99 \times 10^{-2}}{-0.12 + eCOG + trt}$
----	-------	--



CONCLUSIONS

- 1,238 models were generated in <3 minutes.
- The 'best' models generated R² of around 9%, but with significant complexity (level 222).
- A less complex model (complexity level 42) gave R² of 6.5% and AIC of 272.
- SR resulted in lower AIC, higher R² and more accurate mean utility estimates for health states than the usual (linear) modelling approaches often found in HTA submissions.
- The mean utility estimates from AI models can lead to notably different estimates of mean Quality Adjusted Life Years (QALY) and consequent incremental cost-effectiveness Ratios (ICERs).

REFERENCES

1. Khan et al BMJ Open 2015 Jul 2;5(7)
2. SM Lee et al; Lancet Oncol, 2012 Nov;13(11): 1161-70
3. Hussein.S; Genetic Programming in Mathematica <http://www.husseinsspace.com/research/publications/gpinmath.pdf>