Institution: Nottingham Trent University (NTU)		
Unit of Assessment: A03 – Allied Health Professions, Dentistry, Nursing and Pharmacy		
Title of case study: Commercialisation of AI-based algorithms for faster *In silico* drug and diagnostic biomarker discovery.		
Period when the underpinning research was undertaken: 2002 – 2020		
Details of staff conducting the underpinning research from the submitting unit:		
Name: Graham Ball	**Role:** Professor of Bioinformatics	**Period employed by submitting HEI:** 2000 - present
Period when the claimed impact occurred: 1 August 2013 to 31 December 2020		
Is this case study continued from a case study submitted in 2014? No		

1. Summary of the impact

Ball has developed patented Artificial Neural Network based techniques to facilitate extraction of aetiological meaning from complex omics data repositories. NTU spin-out Intelligent OMICS has licenced and commercialised Ball's patented techniques, generating £230k revenues and attracting £2m finance. Intelligent OMICS has helped secure US Defence Threat Reduction Agency contracts with DSTL worth £4.5m to enable rapid diagnosis of infection, particularly trauma induced sepsis, impacting on healthcare in the military. Intelligent OMICS incorporation of AI methods into drug discovery process enabled Cumulus Oncology's successful £1.7m fund raise. Low-cost diagnostic panels have been shown to manage COPD by Mologic, benefitting clinical trial patients, and 56 latent TB cases identified during trials with Wuhan Pulmonary Hospital (China) enabled these patients to receive early treatment.

2. Underpinning research

New approaches for studying disease systems at the genomic, epigenetic, proteomic and metabolomic levels are continually developed to predict responses to therapy or identify new drug targets. The challenges in analysing large molecular datasets lie in their volume, resolution and complexity, and the quality of the data, which can lead to false discovery and overfitting (i.e. falsely elevated diagnostic performance).

Research at NTU, led by Ball over 18 years, has developed a novel Machine Learning systems approach to *in silico* biomarker discovery. Based on Artificial Neural Networks (ANNs), Ball's computational algorithms facilitate analysis of large complex omics datasets, to identify non-linear biomarkers associated with clinical features concordant across multiple datasets. The algorithms study interactions between nodes within a single disease pathway and identify the most influential sets of biomarkers in a given biological system.

The initial concept was to utilise ANNs with a constrained architecture enabling them to discover robust biomarkers from complex omics data. Ball's group was the one of the first to mine mass spectrometry data using an ANN approach; the first study identified biomarker ions that accurately differentiated between two common types of brain tumours, astrocytoma and glioblastoma (**R1**). These ANN algorithms were further developed to model public molecular data repositories like the Gene Expression Omnibus and The Cancer Genome Atlas (**G1, G2, G3, G4**) in order to identify robust molecular diagnostic panels from the whole transcriptome, concordant across multiple data sets. This approach provided a rank order of the molecular biomarkers associated with a given clinical question, overcame the issues of false discovery and overfitting, and improved the predictive performance for unseen cases from disease populations. In a collaborative study (**G5, G6, G7, G8**) (Ball: Joint PI) with Nottingham University Hospitals Trust (NUH), it was used to validate a gene expression profile for the detection of clinical outcomes in breast cancer patients, having reduced a 70-gene signature to just nine genes capable of accurately predicting distant metastases (**R2**). Patents for the ANN data methods were filed in 2009, and a spinout company Intelligent OMICS was launched to exploit the algorithmic IP and the methods (Note: Compandia Ltd was incorporated 12 Feb 2008, then renamed CompanDX Ltd, the company is now CompanDX PVT in China and Intelligent OMICS in the UK).

Further work extended the rank order concept to undertake a meta-analysis over multiple questions and multiple data sets (**G1**). Ball developed a new algorithm that utilised ANN-based network inference (ANNi) methods (**G1, G3**) to model networks of molecular interactions between context-specific biomarkers, for a given pathway or disease, based on whole transcriptomic data. By identifying associations between molecules in a pathway and the remaining molecules in the whole transcriptome, new pathway features for a given disease could be identified and analysed to find the most influential molecular drivers in a given system.

These new approaches were applied in a breakthrough study, with NUH, Universities of Nottingham, Cambridge and Auckland and China's Northeast Normal University, analysing the transcriptomic gene expression levels associated with multiple proliferation features in over 4500 breast cancer cases. Published in *The Lancet Oncology*, the research identified a gene SPAG5 which, was the most influential in the proliferation system and indicates which patients will benefit from chemotherapy, meaning SPAG5 could be a new therapeutic target and a biomarker for the tailoring of breast cancer treatment (**G5, R3**). The study identified a total of 34 biomarkers in the breast cancer proliferation system. These markers had a calculated probability of false discovery of less than $1x10^{-73}$, further demonstrating the efficacy of the ANNi approach. The ANNi methodology was also successful in specifying influential genes that could be new target markers for childhood sarcomas (**R4**); in identifying, via a commercial contract with Syngenta, transcriptomic regulators of ripening in the tomato (**R5**); and in modelling the ATR-CHEK1 system (critical for genomic stability) to predict and validate therapeutic targets for breast cancer (**R6**). Ball's ANN-based technologies have been applied to numerous clinical questions and the results described in 200-plus publications.

3. References to the research

Underpinning research quality is evidenced by rigorously externally peer reviewed outputs:

R1. G . Ball, S . Mian, F . Holding, R.O. Allibone, J. Lowe, S. Ali, G. Li, S. McCardle, I.O. Ellis, C. Creaser, R.C. Rees. An integrated approach utilising artificial neural networks and SELDI mass spectrometry for the classification of human tumours and rapid identification of potential biomarkers. *Bioinformatics* 18 (3), pp. 395-404 (2002) http://doi.org/10.1093/bioinformatics/18.3.395

R2. L.J. Lancashire, D.G. Powe, J.S. Reis-Filho, E. Rakha, C. Lemetre, B. Weigelt, T.M. Abdel-Fatah, A.R. Green, R. Mukta, R. Blamey, E.C. Paish, R.C. Rees, I.O. Ellis, G.R. Ball. A validated gene expression profile for detecting clinical outcome in breast cancer using artificial neural networks. *Breast Cancer Research and Treatment* 120, pp. 83-93 (2010) http://doi.org/10.1007/s10549-009-0378-1

R3. T.M.A. Abdel-Fatah, D. Agarwal, D.-X. Liu, R. Russell, O.M. Rueda, K. Liu, B. Xu, P.M. Moseley, A.R. Green, A.G. Pockley, R.C. Rees, C. Caldas, I.O. Ellis, G.R. Ball, S.Y.T. Chan. SPAG5 as a prognostic biomarker and chemotherapy sensitivity predictor in breast cancer: a retrospective, integrated genomic, transcriptomic, and protein analysis. The Lancet Oncology 17(7), pp. 1004-1018 (2016). http://doi.org/10.1016/S1470-2045(16)00174-1

R4. D.L. Tong, D.J. Boocock, G.K.R. Dhondalay, C. Lemetre, G.R. Ball. Artificial Neural Network Inference (ANNI): A Study on Gene-Gene Interaction for Biomarkers in Childhood Sarcomas. *PLoS ONE* 9(7), Art. No. e102483 (2014). http://doi.org/10.1371/journal.pone.0102483

R5. Y. Pan, G. Bradley, K. Pyke, G.R. Ball, C. Lu, R. Fray, A. Marshall, S. Jayasuta, C. Baxter, R. van Wijk, L. Boyden, R. Cade, N.H. Chapman, P. Faser, C. Hodgman, G.B. Seymour. Network Inference Analysis Identifies an APRR2-Like Gene Linked to Pigment Accumulation in Tomato and Pepper Fruits. *Plant physiology* 161(3), pp/ 1476-1485 (2013) http://doi.org/10.1104/pp.112.212654

R6. T.M.A. Abdel-Fatah, F.K. Middleton, A. Arora, D. Agarwal, T. Chen, P.M. Moseley, C. Perry, R. Doherty, S. Chan, A.R. Green, E. Rakha, G.R. Ball, I.O. Ellis, N. J. Curtin, S. Madhusudan. Untangling the ATR-CHEK1 network for prognostication, prediction and therapeutic target validation in breast cancer. *Molecular Oncology* 9(3), pp. 569-585 (2015) http://doi.org/10.1016/j.molonc.2014.10.013

The high quality of the underpinning research is further indicated by the following funding investments in the research and its dissemination:
G1. Charitable funding for the van Geest Cancer Research Centre. Led by Prof Robert Rees. Funding for algorithm development, data generation, computational costs. 2008 to 2019
G2. Innovate UK. Translation of Bovine Tuberculosis biomarkers to the point of care setting. 101555. Led by Prof Graham Ball as CompanDX Ltd. Total value £980k. Validation and deployment of markers derived from machine learning. Jan 2015 to Jan 2017
G3. Technology Strategy Board Sepsis II Project ANEMONES TSB ref 101191 (reference 15841-108156) Validation of prognostic biomarkers for Sepsis. Led by Karen Kempsall, Health Protection Agency. Total value £1.2M. Development of algorithms for the analysis of molecular data for discovery of biomarkers. July 2012 to Jan 2014
G4. European Union Sixth Framework funded Network of Excellence (Contract number IST-2002-1 508803) BIOPATTERN – Computational intelligence for biopattern analysis in support of eHealthcare. Principal investigator – Dr Graham Ball. Coordinator Prof Emmanuel Ifeachior – School of Computing, Communications and Electronics, University of Plymouth. Total Value €6.2M (Grant value for NTU component €196k/£151k) Jan 2005 to Jan 2008.
G5. NIHR evaluation of the SPAG5 biomarker in the clinical setting. NIHR Co-I Total value £890k II-LA-0417-20004. Deployment of Machine learning derived biomarkers for proliferation.
G6. Breast cancer campaign (Contract Reference 2005nov08) - Transcriptional/Translational Analysis of Breast Cancer Providing A Modern Clinical Classification System. Total Value £174,501 (Grant value for NTU component £18,000). Jan 2006 to Jan 2009
G7. I-Net Nottingham Prognostic Index+ Lead by Prof Ian Ellis, University of Nottingham. Total Value £100,429. Regional funding leading to DPFS grant, below. 2008
G8. MRC DPFS award (DPFSDP003) Nottingham Prognostic Index Plus (NPI+) Lead by Prof Ian Ellis, University of Nottingham. Total Value £767,375. Development of a molecular prognostic index for breast cancer and development of machine learning methods for analysis of molecular data. Aug 2009 to March 2012.

4. Details of the impact

Impacts on commerce and the economy: **NTU's spin-out company Intelligent OMICS Ltd. has commercialised NTU IP and has generated significant revenue and finance investment.**

Successful business Intelligent OMICS Ltd has been built through the commercial exploitation of the ANN methods and ANNi machine learning algorithms developed at NTU for *in silico* biomarker and drug target discovery. The company was rebranded as Intelligent OMICS Ltd in 2019 (see Section 2) to better reflect the key technological breakthrough that its Intuitive Informed Intelligence (I3) machine learning platform, built around the NTU research, could now process 50 million models to a given dataset per hour. This innovation has shortened analysis, that would have previously taken six to seven months, down to two to three weeks. The IP assigned to Intelligent OMICS from NTU includes the patent "Data analysis method and system" (US8788444 B2, and equivalents) covering Ball's machine learning methodologies. The company has attracted £148,000 from Innovate UK (total grant £873,000, grant reference 101555, Nov 2013 to Jan 2017). Over the current REF2021 impact period, the company "*has generated over £230,000 of revenue from third parties – pharma and biotech companies, and UK healthcare-related institutions*" (**S1**), created 4 jobs, and the company has "*financed and completed over £5,070,000 of research and clinical trials in China"* (**S1**). This latter figure, of which £2m pro-rata has been invested during the period 1 August 2013 to 31 July 2020 (latest information available within REF2021), includes investment over 5 years from Hubei provincial government, China.

***Impacts on commerce and the economy*: NTU and Intelligent OMICS' research has stimulated foreign direct investment of £4.5m from United States DTRA for UK Dstl-led consortia.**

Defence Science and Technology Laboratory (Dstl, an executive agency, sponsored by the UK Ministry of Defence), its military stakeholders, and NTU/Intelligent OMICS Ltd. have received contracts from the US Defence Threat Reduction Agency (DTRA) based on the application of NTU and Intelligent OMICS Ltd. machine learning algorithms. The arising IP is being jointly exploited

for a civilian context by Ploughshare Innovations, Dstl's commercial arm, and Intelligent OMICS. The letter of corroboration from a Senior Principal Statistician in the Cyber and Information Division of Dstl states, "*The research collaboration is focussed on bringing better health care to the military and Prof Graham Ball's research output is having impact on the defence community through rapid diagnosis of infection, particularly trauma induced sepsis*" (**S2**), and furthermore, "*the civilian applications could really assist with the fight against antimicrobial resistance (AMR). The methods that have come from the collaboration have identified previously unseen patterns in gene behaviour in the presence of infection and have helped secure international contracts totalling £4.5m in US funding. Contract references DSTLX 1000896368 AM 2, HDTRA1-14-24-FRCWMD-BAA. A further project PREDICTS HDTRA118-CBA-01-1-0044 is being negotiated (Value £2.1M)*" (**S2**).

***Impacts on commerce and the economy*: commercial adoption of NTU spin-out Intelligent OMICS' algorithms have provided Cumulus Oncology with a significantly more cost-effective method to target drugs for development, and facilitated their successful £1.7m fund raise.**

Using public and private molecular data Ball through Intelligent OMICS has brought about change in practice and generated commercial income through a pathway-based data mining approach based on Artificial Neural Network Inference ANNi (**R4**) algorithmic process that identifies disease/pathway-specific molecular drivers. This approach replaces the expensive and time-consuming drug screening process with the modelling of biological pathways in a disease. These highly influential molecules operating in a pathway make excellent biological targets for drugs. As a result, Intelligent OMICS has entered into a strategic partnership with Cumulus Oncology. Work with Cumulus has provided insights into key pathways of interest to the company around DNA repair (DDR and CHEK1) in lung cancer. This has enabled Cumulus to define further druggable targets of interest and develop biomarker hypotheses for the drug assets under development, facilitating a recent investment into Cumulus from Private Investors and Scottish Enterprise. A key requirement of this investment was the incorporation of AI methods into the drug discovery process. The CEO of Cumulus Oncology stated, "*Through our collaboration with Intelligent OMICS, the application of the AI based methods developed by the company has provided Cumulus with insights into key pathways of interest for current and future assets. In turn, this has facilitated a recent successful fund raise of £1.7m by Cumulus Oncology*" (**S3**).

***Impacts on commerce and the economy:* commercial adoption of NTU spin-out Intelligent OMICS' algorithms within Mologic's new low cost rapid diagnostic tests.**

The ANN algorithms developed by Ball have been offered as a service to diagnostics companies, where they are used to rapidly discover optimised panels of biomarkers, incorporate the biomarker panel into a diagnostic model and deliver that model as a piece of software for incorporation onto a diagnostic device. The Chief Scientific Officer of Mologic Ltd, which employs 70 people, stated that, "*With the help of the machine learning based techniques of Prof Graham Ball and the spin out company Intelligent OMICS Ltd, our company is now able to identify and develop optimised biomarker panels incorporated into point of care diagnostic devices*" (**S4**), and furthermore, that "*The methods and diagnostic interpretation principles coming from Professor Ball's research have greatly improved the efficiency, sensitivity and specificity of the ground-breaking diagnostic tests we are developing*" (**S4**). In November 2019, Mologic announced it had started a clinical trial (COPE-WEL; 263 patients) to validate its patented, urine-based, multi-biomarker self-test Headstart® for the early detection of exacerbations in patients living with moderate to severe Chronic Obstructive Pulmonary Disease (COPD) (**S4**, **S5**). Clinical decisions are being made within the COPE-WEL trial based on the early detection of COPD exacerbation, benefiting patients by providing timely and appropriate interventions. The test can predict the early onset of a COPD exacerbation up to six days before symptom onset, or up to nine days before clinical diagnosis by a healthcare provider. It "*depends on an associated algorithm developed by Professor Ball*" (**S4**). Following proven success from the validation trials, this urine test will be available for launch in 2021. Another complex biomarker test "*made simple by Prof Ball's technology"* (**S4**) has been developed for sepsis, and in September 2019 Mologic announced it had finished recruiting 550 people to a clinical trial being carried out at University College London Hospitals (**S6**). Clinical

decisions separating bacterial infections from inflammatory responses where there is no bacterial cause have help managed the way patients are treated leading to patient benefit.

Impact on human health: **trial of a new clinical test for latent tuberculosis in collaboration between NTU spin-out Intelligent OMICS and Wuhan Pulmonary Hospital (China), successfully detected latent cases of the disease enabling patients to be treated, and led to phase 2A trials which have been completed.**

Ball, through Intelligent OMICS Ltd., has developed and applied algorithms to generate a panel of diagnostic biomarkers which identify the presence of Latent Tuberculosis in otherwise healthy individuals. These inventive methods were developed by Ball and patent protected (GB201211158A, priority date 22 June 2012, equivalents EP2864499 B1, US2015197806 A1, WO2013190321 A1, "Biomarkers for determining the M. tuberculosis infection status", Inventor G.R. Ball). In 2017 Intelligent OMICS Ltd. embarked on clinical evaluation trials in collaboration with Wuhan Pulmonary Hospital (WPH) in Hubei Province, China to exploit the predictive capability of Ball's patented technology (**S7**). Initial investment to run the trial was received from the Hubei provincial government, followed within the current REF period by a further 3551 Foreign talent programme grant of £112k (¥1.0m) from Wuhan City and Wuhan Biolake development fund (**S8**), plus additional match funding investment from Wuhan Pulmonary Hospital. This project received the backing of Wuhan's Vice Mayor and the province Consul, who were present at a contract signing ceremony (**S9**). The phase 1 trial revealed 6 asymptomatic cases of Latent Tuberculosis, and the phase 2A revealed 50 latent cases, via simple blood test, enabling these patients to receive early treatment. The results of the completed phase 2A Clinical trial are currently being decoded. Wuhan City £224k (¥2.0m) funding is already in place to commence a Phase 2B trial that will examine 250 patients.

5. Sources to corroborate the impact (* participant in the process of impact delivery)

S1*. Testimonial letter: Chief Executive Officer, Intelligent OMICS Ltd.
S2*. Testimonial letter: Senior Principal Statistician in the Cyber and Information Division, Defence Science Technology Laboratory.
S3*. Testimonial letter: Founder and Chief Executive Officer, Cumulus Oncology Ltd.
S4*. Testimonial letter: Chief Scientific Officer, Mologic Ltd
S5. Web-link: Mologic Ltd. press release on the COPD clinical trial: https://mologic.co.uk/mologic-launches-clinical-trial-for-validation-of-point-of-care-copd-exacerbation-alert-system/
S6. Web-link: Mologic Ltd. press release on the sepsis clinical trial: https://mologic.co.uk/mologic-developing-sepsis-test-using-multimarker-lateral-flow-technology/
S7*. Confidential document: Collaboration agreement with Wuhan Pulmonary Hospital and CompanDX Ltd (now Intelligent OMICS Ltd).
S8. Confidential document: Wuhan East Lake High-tech District, funding for "Diagnostic techniques of Tuberculosis" project, application and receipt for RMB 500k
S9. Confidential document: Contract signing ceremony, Wuhan, China