Evaluation of Participants' Reaction and Learning in a Taught Analytics and Modelling Academy Program in U.K.'s National Health Service

Daniel Gartner
School of Mathematics
Cardiff University and
National Health Service, Wales
Cardiff, United Kingdom
gartnerd@cardiff.ac.uk

Tracey J. England
School of Heath Sciences
University of Southampton
Southampton, United Kingdom
t.j.england@soton.ac.uk

Doris A. Behrens
Department for Economy and Health
Danube University Krems
Krems, Austria
doris.behrens@donau-uni.ac.at

Izabela Spernaes
Mathematical Modelling Unit
Aneurin Bevan University Health Board
National Health Service, Wales
Newport, United Kingdom
izabela.spernaes@wales.nhs.uk

Joanne Buchanan
Mathematical Modelling Unit
Aneurin Bevan University Health Board
National Health Service, Wales
Newport, United Kingdom
joanne.buchanan@wales.nhs.uk

Paul R. Harper
School of Mathematics
Cardiff University
Cardiff, United Kingdom
harper@cardiff.ac.uk

***Abstract*—Recent research has highlighted the need to invest in the development of healthcare analytics capability. However, the contents of such programs and how they should be delivered to maximize the learning outcome are unclear. In this paper, we provide insights into the learning within the first two cohorts of modelling fellows successfully trained in an analytics and modelling academy run within the National Health Service (NHS) Wales, U.K. The participants followed a taught healthcare analytics and mathematical modelling program tailored for senior staff members including managers and clinicians. We build our learning evaluation framework on Kirkpatrick's training evaluation model and participants filled in questionnaires with respect to their level 1 (reaction) and level 2 (learning) experience after each module. In addition, we asked the participants about their self-assessments during three time points in the program. The qualitative feedback results revealed that the participants appreciate the learning and reflect where they could use the new developed skills in practice. They also provided useful suggestions for improving the program. The participants' aggregated quantitative self-assessments show a statistically significant increase in competence. In conclusion, this may lead to a behavior change in applying the methods on the job (level 3) and, ultimately, improve level 4 outcomes through analytics-driven healthcare improvement.**

***Index Terms*—Computer and Information Science Education, Queuing Theory, Simulation, Modeling Methodologies, Modeling and Prediction, Healthcare, Optimization of Service Systems**

I. Introduction and Motivation

Data on the medical care of hundreds of millions of patients are piling up in electronic health records in clinics and hospitals around the world [1]. While timely and effective Operations Management (OM) methods have been used in the public sector for decades [2], the history of using data analytics (e.g. clinical pathway mining) and OM methods (e.g. operating room scheduling) in healthcare primarily gained attention in or after the year 2000 and since then has grown exponentially [3], [4].

There has been recognition of the role that mathematical modelling and Operations Research (OR) have to play in decision making across the U.K.'s National Health Service (NHS) over the last 20 years: The inauguration of the National Institute for Health and Care Excellence (NICE) can be seen as an example of responding to pressures of tight public sector fiscal settlements and for using innovative approaches. The goal is to ensure excellence by applying OR methods such as Computer Simulation with the ultimate objective to improve outcomes for patients and streamlining services.

Previously, Analytics and Operations Research in the NHS has been the domain of specialists often operating from university centres or analysis units in a supporting role. In NHS Wales, there now exists almost ten years worth of experience in the Aneurin Bevan University Health Board, a part of the NHS, where researchers in residence team up with senior managers and clinicians. The setting strongly suggests that if the service is to get the best from these approaches, further effort has to be made to prepare, teach and support the NHS workforce to understand and use data analysis and OM/OR approaches to support decision making. Responding to this need, mathematical modelling in that Health Board is evolving. They develop, run and evaluate a concise program for teaching healthcare analytics and OM excellence to a small but high potential cohort of service staff. The program consists of two parts: stage one includes the taught part including courses in

The authors thank The Health Foundation for their financial support in the evaluation of the program.

healthcare analytics and modelling. The second stage involves the successful completion of a project.

Many programs do not go beyond evaluation of a single learning event which motivated us to not only evaluate single courses but also the learning in terms of methods across different courses. Our short-term goal in terms of evaluation is to find out the participants' reaction and learning while in the long term we would evaluate behaviour change and improving quality of care by using our methods We evaluated the two cohorts using Kirkpatrick's first two levels of the 4-level framework [5] shown in Figure 1. This training evaluation model has been used widely throughout education.

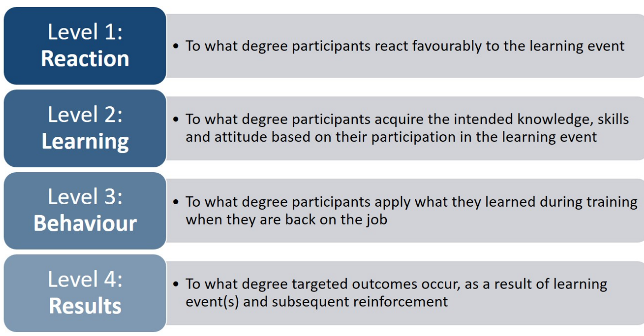

Fig. 1: Kirkpatrick's Model for Learning Evaluation

For cohort one and two, questionnaires to gather participants' level 1 and level 2 feedback were distributed. Additionally, for the cohort two evaluation, structured self-assessment questionnaires were handed out to the participants at the beginning of the program, half-way through and at the end of the program. Our quantitative self-assessment results reveal that the participants have a steep learning curve from the start until the end of the program. This is true for almost every taught topic. The qualitative feedback results revealed that the participants appreciate the learning and reflect where they could use the new developed skills in the NHS. Adopting the methods learned in the program in the day-to-day work environment may lead to a significant behavior change and improving healthcare services.

The remainder of this paper is structured as follows: In section II, we will provide an overview of related work and initiatives that focus on the development and evaluation of similar courses, modules and programs. Section III provides an overview of the material and methods that we used in evaluating our program. Section IV shows summary statistics and results of the participants' learning experience. Section VI closes the paper with conclusions.

II. Related Work and Programs

Award-winning analytics and operations management programs have been established for many years in universities such as Massachusetts Institute of Technology [6] and Carnegie Mellon University [7]. More specifically, the teaching of Operations Research using case studies has been reviewed by Drake (2019) [8] who concludes that there is a shift in OR teaching from technical mathematics to model building and application. Although there is a large need in the planning of effective and efficient delivery of healthcare services, only a few tailored healthcare analytics programs [9] or healthcare modules in e.g. MBA programs exist or have been evaluated by learning frameworks. We will summarize existing programs below.

Kopcso et al. (2016) [10] provide a case study and teaching notes in the area of pharmaceutical development [11]. The case study is aimed at undergraduate engineering and graduate MBA students. The particulars of the study involve a decision problem in which a firm must select a development option for two potential drugs used to treat idiopathic pulmonary fibrosis. The decision problem involves two players – one a larger firm with experience in clinical trials and the second a smaller firm holding the license for a particular drug of interest. The case illustrates the importance of structured decision making and supports the use of decision trees for determining the potential costs and benefits of various courses of action.

Proano (2016) [12] uses a case study approach to teach production management to engineering undergraduate students [11]. The case describes planning for and dealing with the outbreak of pandemic influenza on a university campus. In this case, which is designed to run over a term, students are responsible for developing mechanisms for predicting the number of individuals who will become infected and for identifying what policies should be implemented to deal with the outbreak once it becomes pandemic. The case introduces a number of technical components, such as forecasting and epidemiological disease models, and places them in a decision context in which the future may not be predictable from the past.

Apart from the taught case studies published in the literature, programs exist that teach analytics capability in healthcare such as the Health Service Modelling Associates Programme [13] carried out at the University of Exeter. The major difference to the program evaluated in this paper is, however, that the focus is on the delivery of multiple mathematical modelling and programming methods such as discrete optimization, simulation, system dynamics, scheduling, and Visual Basic for Applications.

III. Material and Methods

Tables I and II provide an overview of the first and the second cohort when the program was run, respectively. As can be seen, the program is front-loaded with courses in the first months followed by individual project coaching sessions.

The figures also reveal that more courses were run in the second cohort of the program as compared to the first one with the intersection of the following (core) modules:

- Geographical Analytics
- Forecasting
- Statistics
- Queueing
- Staffing / Rostering and
- Scheduling

TABLE I: The analytics and modelling academy program run for the first cohort

Month	Activity	Duration
1	Classroom based teaching 1: • Introduction to the programme • Overview of projects that have been selected • Geographical Analytics	1 day
	Classroom based teaching 2: • Forecasting	1 day
2	Classroom based teaching 3: • Essential Statistics in Healthcare II (Inferential Statistics)	1 day
	Classroom based teaching 4: • Queueing in Healthcare	1 day
3	Classroom based teaching 5: • Staffing, Workforce Planning and Rostering	1 day
	Classroom based teaching 6: • Scheduling in Healthcare	1 day
	Classroom based teaching 7: • Your project and the data you need	½ day
4–8	Weekly mandated coaching sessions	TBA with the supervisor
10	Celebration event: • Presentation of project work	½ day

TABLE II: The analytics and modelling academy program run for the second cohort

Month	Activity	Duration
1	Classroom based teaching 1: • Introduction to the programme • Overview of projects that have been selected • Geographical Analytics	1 day
1	Classroom based teaching 2: • Forecasting	1 day
1	Tutorial Session 1	½ day
2	Classroom based teaching 3: • Systems Thinking	1 day
2	Classroom based teaching 4: • Queueing and Simulation in Healthcare	1 day
2	Tutorial Session 2	½ day
3	Classroom based teaching 5: • Statistics	1 day
3	Classroom based teaching 6: • Rostering	1 day
3	Classroom based teaching 7: • Scheduling	1 day
4	Classroom based teaching 8: • Visual Basic for Applications	1 day
4	Tutorial Session 3, Project Introduction	1 day
5–11	Weekly mandated coaching sessions	
12	Celebration event: • Presentation of project work	½ day

Cohort one consisted of seven participants and cohort two had six participants which results in a sample size of $n = 13$ participants for the level 1 and level 2 evaluation. Participants had a variety of different backgrounds and their job roles included:

- Head of Business and Financial Planning
- Project Manager
- Head of Strategic Finance & Innovation
- Clinical Fellows
- Business and Performance Manager
- Program Manager
- Senior Business Analyst
- Information Development Manager
- Theatre Performance Officer
- Assistant Directorate Manager

A. *Level 1 evaluation*

For the level 1 evaluation in cohort one and two, participants were asked to fill out a feedback questionnaire after the end of each course which is provided in Figure 8 in the Appendix.

B. *Level 2 evaluation*

For the level 2 evaluation, participants filled out a feedback questionnaire after the end of each course which is provided in Figure 9 in the Appendix. We used word-stemming and removed stop words which is a common technique in free-text analytics. The feedback was then put into a word-cloud processor in order to gather themes as Section IV-B will reveal.

C. *Self-assessment evaluation*

Additionally to the level 1 and level 2 evaluation, cohort two participants were asked to indicate their current level of knowledge at the start of the program, mid-way through the program and at the end of the program. The 25 analytics domains are provided in Figures 8–9 in the Appendix. We evaluated the feedback using a 25 point scale reaching from Information ("Knowing what the tool is") up to Wisdom ("Can teach theory and use the method").

IV. Results

In the following sections, we will provide details of the evaluation results.

A. *Level 1 evaluation*

The level 1 evaluation (reaction) results are shown in Figure 2. The boxplots reveal that in all but one question ("Question 6: I will be able to apply the knowledge and skills I have learned from today", see Figure 6 in the Appendix, the median evaluation was 9 (strongly agree). For this particular question 6, the standard deviation of responses was highest ($\sigma = 1.979$). This suggests that for some participants, coaching will be needed to apply the techniques in the practical setting which is why the program has mandatory coaching sessions in the project phase, see Figure II.

B. *Level 2 evaluation*

In this evaluation, a questionnaire (see Figure 7 in the Appendix) was used to ask the participants to rank the two questions from 0 (strongly disagree) to 10 (strongly agree) after each of the courses:

1) Will the knowledge and skills that you have gained today help you to improve the way you work?
2) How confident are you that you will be able to apply what you have learned in the workplace?

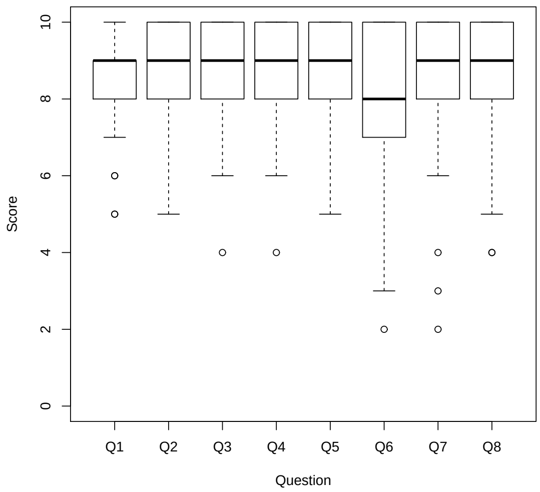

Fig. 2: Level 1 evaluation results with 0 meaning 'strongly disagree' and 10 meaning 'strongly agree'

For the first question, the scores varied from 4 to 10 across different days with the average of 8.9 and a median of 9. The vast majority (93%) of all responses scored 8 or above meaning that the participants strongly feel that the knowledge and skills will improve the way they work. Furthermore, the analysis of the second question showed that the scores varied from 2 to 10 across different days with the average of 7.8 and a median of 8. The majority (70%) of all responses scored 8 or above meaning that the participants are strongly confident to be able to apply what they have learned in the workplace. Figure 3 shows the themes identified in the open questions of cohort 1 and 2 level 2 evaluation. The figure reveals that participants appreciate concepts like correlation in the courses but also rostering are useful topics in the planning of services.

Fig. 3: Themes in the level 2 evaluation

C. Self-assessment evaluation

We compared the self-assessments' aggregate numbers with respect to pre-, mid- and post program. The results are shown in Figure 4. The boxplots reveal that the median self-evaluation score strictly monotonically increases. A more detailed analysis revealed that the median values are 2, 12 and 16 for the pre-, mid- and post-self-evaluation, respectively. This suggests a less dramatic increase in the self-evaluation score between mid and post as compared to pre and mid assessment.

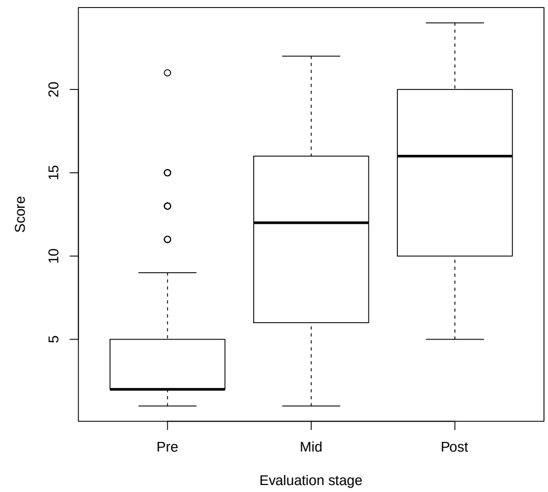

Fig. 4: Box plots for the comparison of the pre-, mid- and post-program self-assessments

1) Significance tests: We carried out a statistical analysis of the three evaluation milestones. Using the Shapiro-Wilk test implemented in R [14], we first tested whether the respondents' responses in each of the evaluations follow the normal distribution which was rejected ($p < 0.001$) in all cases. We then carried out a Wilcoxon test for differences in the respondents' learning of the pre- and mid-evaluation. The result was that there was a significant increase in the self-assessment score ($p < 0.001$) between the pre- and mid-evaluation. Furthermore, we found a significant increase in the self-assessment score ($p < 0.001$) between the mid- and post-evaluation.

2) Module-Specific Learning: Figure 5 provides a radar chart from the 2nd cohort.

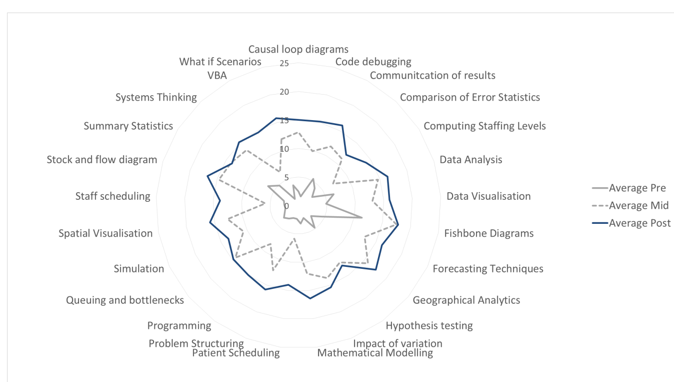

Fig. 5: Radar chart for the self-evaluation of cohort 2 ($n = 6$)

The data is broken down by pre-, mid- and post-program self- assessment. A closer investigation of the module-specific learning revealed that some modules have a higher difference in the mean pre-, mid- and post evaluations as compared to others. We tested whether there was significant learning in each of the specific taught methods between the pre- and the post-evaluation. The figures provided in Table III reveal that there is a statistically significant difference between the pre- and the post evaluation score.

TABLE III: Results of the comparison between the pre- and the post-evaluation (*: One-tailed Wilcoxon test ($\alpha = 5\%$, $p = 0.0217$))

Method	Avg. Pre-Score	Avg. Post-Score	Avg. Difference
Causal loop diagrams	2.5	15.0	12.5*
Code debugging	1.7	15.2	13.5*
Communication of results	5.3	16.0	10.7*
Comparison of Error Statistics	4.2	12.2	8.0*
Computing Staffing Levels	2.8	14.0	11.2*
Data Analysis	6.5	16.4	9.9*
Data Visualisation	4.8	16.0	11.2*
Fishbone Diagrams	11.3	17.8	6.5*
Forecasting Techniques	4.5	16.2	11.7*
Geographical Analytics	2.8	17.6	14.8*
Hypothesis testing for significance	4.8	13.0	8.2*
Impact of variation	2.3	15.4	13.1*
Mathematical Modelling	3.2	16.4	13.2*
Patient Scheduling	2.3	14.0	11.7*
Problem Structuring	2.3	15.8	13.5*
Programming and loop structures	2.8	15.0	12.2*
Simulation	2.7	13.6	10.9*
Spatial Visualisation	2.3	15.8	13.5*
Staff scheduling	2.5	13.8	11.3*
Stock and flow diagram	2.7	16.8	14.1*
Summary Statistics	6.3	13.8	7.5*
Systems Thinking	4.8	15.2	10.4*
Understanding of Queuing and bottlenecks	3.3	14.8	11.5*
VBA	1.3	14.6	13.3*
What if Scenarios	3.7	15.8	12.1*

V. Limitations

The sample size in our study was $n = 13$ participants. Although our pre- and post-evaluation revealed a significant increase in terms of learning, there is still a 2.17% chance that we falsely reject the null-hypothesis that there is no difference between the pre- and post-score. We would hope that this likelihood decreases with an increasing n. Another limitation on the generalizability is that the program does not consist of modules such as Artificial Intelligence or Machine Learning. We therefore cannot conclude whether teaching these methods to healthcare managers and clinicians would be useful. Another limitation is that we only focused on Kirkpatrick levels 1 and 2 but not on the behaviour change (level 3) and how the taught methods influence outcomes (level 4).

VI. Conclusions

In this paper, we have developed and applied a learning evaluation framework for reaction and learning in a healthcare analytics and modelling program in U.K.'s National Health Service (NHS). The results of Kirkpatrick's level 1 revealed a positive reaction to the course content, learning environment and that the program contributes to the participants future development. The level 2 evaluation showed that the participants appreciated the learning of some analytics topics. Finally, in the self-assessment evaluation we observed significant learning in all 25 analytics topics that are taught. In conclusion, the teaching may lead to a behavior change in applying the methods in day-to-day practice which could be part of a follow-up evaluation in future work. Finally, our developed framework can be used to evaluate other taught programs in healthcare analytics and operations management.

References

[1] R. Winslow. Medical records data offers doctors hope of better patient care. *The Wall Street Journal*, 2021. last accessed on 12 March 2022, https://www.wsj.com/articles/medical-records-data-offers-doctors-hope-of-better-patient-care-11638543613.

[2] S.M. Pollock, M.H. Rothkopf, and A. Barnett. *Operations research in the public sector: An introduction and a brief history*. Elsevier, 1994.

[3] B. Cardoen, E. Demeulemeester, and J. Beliën. Operating room planning and scheduling: A literature review. *European Journal of Operational Research*, 201(3):921–932, 2010.

[4] E. Aspland, D. Gartner, and P. Harper. Clinical pathway modelling: a literature review. *Health Systems*, 10(1):1–23, 2021.

[5] D.L. Kirkpatrick. *Evaluating training programs*. Tata McGraw-Hill Education, 1975.

[6] G. Perakis and D. Rosenfield. The MIT Leaders for Global Operations Program. *Interfaces*, 48(3):189–203, 2018.

[7] R. Krishnan, A. Blumstein, and J. Nehlsen. Designing and Sustaining an Award-Winning Analytics Educational Program: The Heinz College of Information Systems and Public Policy at Carnegie Mellon University. *Interfaces*, 48(4):372–383, 2018.

[8] M.J. Drake. Teaching OR/MS with Cases: A Review and New Suggestions. *INFORMS Transactions on Education*, 19(2):57–66, 2019.

[9] R. Parks and F. Cobb Payton. Designing a healthcare data analytics course: A contextual active learning approach. *Proceedings of the 2018 American Conference of Information Systems*, 2018.

[10] D.P. Kopcso, H. Simon, and A. Gao. Case–idiopathic pulmonary fibrosis. *INFORMS Transactions on Education*, 16(3):107–109, 2016.

[11] J. Blake. Introduction to the Special Issue: Teaching healthcare operations research, Part 1. *INFORMS Transactions on Education*, 16(3):85–86, 2016.

[12] R.A. Proano. Case article—dealing with the bug in the classrooms: Planning for a pandemic. *INFORMS Transactions on Education*, 16(3):93–96, 2016.

[13] M. Pitt and K. Stein. PenCHORD: Helping The NHS Make Informed Decisions Using Operational Research. *Impact*, 2018(1):22–25, 2018.

[14] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2014.

APPENDIX

ABCi
Capability | Innovation | Delivery

GIG CYMRU NHS WALES | Bwrdd Iechyd Prifysgol Aneurin Bevan University Health Board

ABC*i* - Participant Evaluation

Name: ______ **Date:** ______

Course: ______ **Venue:** ______

Instructions:

Thinking about the Session you have just attended, please indicate to what degree you agree with each statement circling one number using the rating scale 0 - 10:

e.g. 0 1 2 3 4 5 6 7 8 9 10

Level 1 – Reaction

	Strongly Disagree										Strongly Agree
The room environment helped me to learn	0	1	2	3	4	5	6	7	8	9	10
There were no major distractions that interfered with my learning	0	1	2	3	4	5	6	7	8	9	10
I engaged with what was going on during the session	0	1	2	3	4	5	6	7	8	9	10
The exercises and activities undertaken aided my learning	0	1	2	3	4	5	6	7	8	9	10
I was given adequate opportunity to practice what I was learning	0	1	2	3	4	5	6	7	8	9	10
I will be able to apply the knowledge and skills I have learnt from today	0	1	2	3	4	5	6	7	8	9	10
The content of the session will help contribute to my future development	0	1	2	3	4	5	6	7	8	9	10
I would recommend the session to my co-workers	0	1	2	3	4	5	6	7	8	9	10

Fig. 6: Kirkpatrick level 1 questionnaire

ABCi
Capability | Innovation | Delivery

GIG CYMRU NHS WALES | Bwrdd Iechyd Prifysgol Aneurin Bevan University Health Board

Level 2 - Learning

Will the knowledge and skills that you have gained today help you to improve the way you work?

Strongly Disagree	0	1	2	3	4	5	6	7	8	9	10	Strongly Agree

Thinking about what you learned today, what will you be able to apply into your role?

How often do you think you will be able to apply your new skills into your role?

Not at all	Monthly	Weekly	Daily

How confident are you that you will be able to apply what you have learned in the workplace? (Circle one rating)

Not at all confident	0	1	2	3	4	5	6	7	8	9	10	Extremely confident

What assistance or resources will you need to successfully apply what you have learned back in the workplace?

Additional Feedback

Please share any additional comments, such as material that was particularly helpful or material you would have liked to have seen included.

Fig. 7: Kirkpatrick level 2 questionnaire

ABCi

GIG CYMRU NHS WALES | Bwrdd Iechyd Prifysgol Aneurin Bevan University Health Board

Analytics and Modelling Academy – Self Assessment

Name: ______ Date: ______

Instructions

Please indicate your current level of knowledge by shading the square on the scale from 1 - 25

E.g.

Tools, methods or skills	Information: Know what the tool is (1 - 5)	Skill: Can apply in identified situations (6 - 10)	Knowledge: Know how, when and where to use (11 - 15)	Understanding: Have experience of, can adapt and explain why (16 - 20)	Wisdom: Can teach the ory and use the method (21 - 25)
Data Analysis					
Mathematical Modelling					
What if scenarios					
Geographical Analytics ie. shortest distance computations					
Spatial visualisation					
Forecasting techniques					
Comparison of error statistics					
Systems Thinking					
Causal loop diagrams					
Fishbone Diagram					
Stock and flow diagram					
Summary Statistics					

Fig. 8: Part 1 of the 25 point self-assessment questionnaire

ABCi

GIG CYMRU NHS WALES | Bwrdd Iechyd Prifysgol Aneurin Bevan University Health Board

Tools, methods or skills	Information: Know what the tool is (1 - 5)	Skill: Can apply in identified situations (6 - 10)	Knowledge: Know how, when and where to use (11 - 15)	Understanding: Have experience of, can adapt and explain why (16 - 20)	Wisdom: Can teach the ory and use the method (21 - 25)
Hypothesis testing for significance (e.g. T-test)					
Visual Basic for Applications (VBA)					
Programming and loop structures (if... and for...)					
Code debugging					
Impact of Variation					
Understanding of Queueing and bottlenecks					
Simulation					
Computing Staffing levels					
Patient Scheduling					
Staff Scheduling					
Communication of results					
Problem structuring					
Data Visualisation					

Fig. 9: Part 2 of the 25 point self-assessment questionnaire