KEYNOTE Reimagining & Accelerating Health Equity, Access & Diversity

DR MARC FARR

Chief Analytical Officer

East Kent Hospitals University NHS Foundation Trust

Re-imagining Health Equity

Prof. Marc Farr
Chief Analytical Officer,

East Kent Hospitals NHSFT
Kent and Medway ICB

NHS
East Kent
Hospitals University
NHS Foundation Trust

Re-imagining Health Equity

The background to analysis of inequalities

Analysis techniques that are now available

Imagining the future – the opportunity and the challenge

Marc

Chief Analytical Officer

Ask me:

Are we providing fair and equitable care at our hospital?

What are the regional plans for Population Health Management to address inequalities?

How are we collaborating with the local universities on recruitment and research?

NHS
East Kent
Hospitals University
NHS Foundation Trust

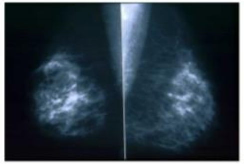
It was 20 years ago today…

Did you know?

…for every tube stop east of Westminster, life expectancy drops by a year…

Did you know?

…women diagnosed with breast cancer in their 40s from an affluent background are more likely to have a lumpectomy than a mastectomy …



Did you know?

…there are as many unplanned pregnancies among the over 40s as there are for the under 20s…

Did you know?

…men diagnosed with bowel cancer in their 50s from a poor background are more likely to have a colostomy bag than a bowel reconstruction…

NHS
East Kent Hospitals University
NHS Foundation Trust

- There are now opportunities to use data and analysis to provide assurance around equitable access, patient experience and outcomes for services delivered at acute hospitals and the importance of putting in place a governance framework for monitoring and tackling any health inequalities that might exist.

- Key points:

    - For the first time, there is national guidance that NHS organisations must be able to disaggregate their board reports by ethnicity and deprivation, in response to national concerns around differences in risk and vaccination uptake for Covid, and waiting times in the elective backlog. Guidance also requires executive leadership in this area.

    - Initial exploratory analysis conducted internally at EKHUFT, including maternity outcomes and likelihood of breaching 18 weeks on elective waiting lists, suggests some inequalities exist between different demographic groups.

    - An informal working group has been established between the Information team and two of our clinicians, seeking to ensure clinical input in understanding causes of health inequalities.

    - A more formal approach to governance is recommended. Information team plan to consider patient access, patient experience and patient outcomes - given that access is currently a key topic, we recommend that inequalities assurance is the responsibility of the COO.

Today…

EKHUFT: Approach to Inequalities analysis

Phase 1 – Q 1 22/23
Bronze

- Counts and Rates
- Crude data
- Examples include:
  - Staffing sexual orientation analysis
- Maturity level:
  - Pie Chart

Phase 2 – Q2 22/23
Silver

- Benchmarking versus all activity including statistical tests.
- Crude data
- Examples include:
  - Paediatric DNA rates
  - Continuity of Care for Maternity
  - Elective waiting list breakdown
  - Rates of falls
- Maturity level:
  - SPC

Phase 3 – Q4 22/23
Gold

- Regression and data mining
- Crude data
- Examples will include:
  - Factor analysis for DNA rates
- Maturity level:
  - Regression/Forecast

Phase 4 – Q2 23/24
Platinum

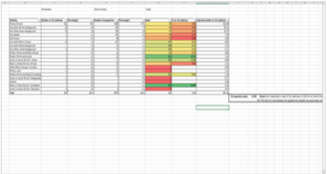
- Standardised rates
- Adjusted data
- Examples will include:
  - Overall equality and variation measure
- Maturity level:
  - HSMR

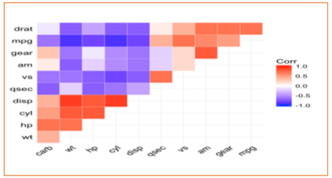
Statistical Tools

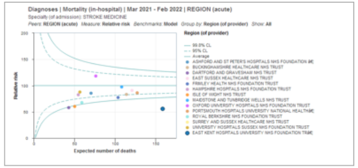
Segmentation Tools

r/Python

Acorn or Mosaic, Look Ups and Stratification Tools







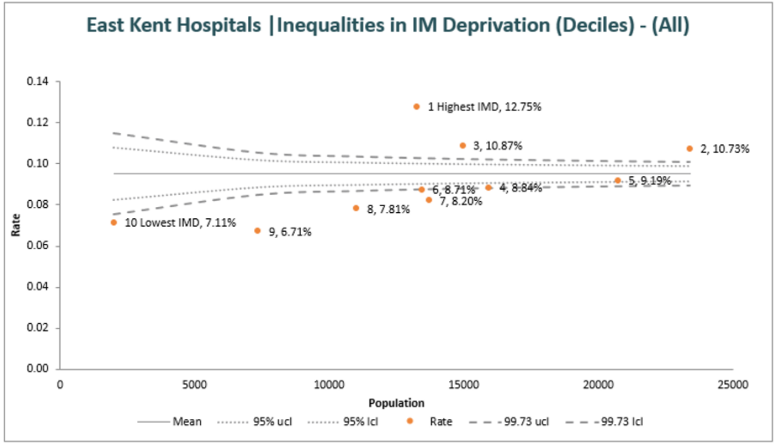
East Kent Hospitals University NHS Foundation Trust

Paediatric DNA rates

The higher the level of deprivation the more likely families with children are to DNA.

Unknown is removed

Row Labels	Sum of Num	Sum of Denom	Index	Expected	Rate
1 Highest IMD	1,692	13,273	134	1263.7	12.7%
2	2,512	23,401	113	2228.0	10.7%
3	1,626	14,962	114	1424.5	10.9%
4	1,408	15,936	93	1517.2	8.8%
5	1,904	20,728	96	1973.5	9.2%
6	1,171	13,446	91	1280.2	8.7%
7	1,124	13,710	86	1305.3	8.2%
8	860	11,018	82	1049.0	7.8%
9	494	7,366	70	701.3	6.7%
10 Lowest IMD	143	2,010	75	191.4	7.1%



Model Training and Evaluation

- Type of problem is Classification

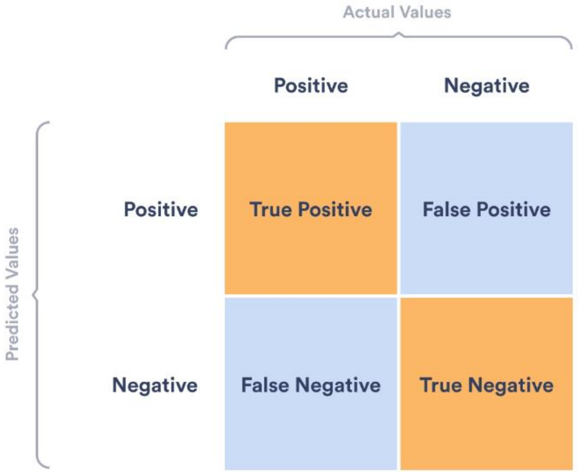
Model	Example Use Cases	Concept in Detail That Produces the Output
Logistic Regression	Predicting the likelihood of an event (e.g., churn prediction, spam detection).	It uses the logistic function to model a binary dependent variable. The model learns weights for the features that maximize the likelihood of reproducing the observed outcomes.
Decision Tree	Decisions that require a clear interpretation and understanding of the model (e.g., medical diagnosis).	It uses a tree-like model of decisions based on the features. Each node in the tree splits the data based on a feature value, and the final decision is based on the leaf nodes.
Random Forest	General classification and regression tasks that require robustness to noise and outliers (e.g., predicting customer satisfaction).	It uses an ensemble of decision trees, each trained on a different subset of the data. The final output is the mode (for classification) or mean (for regression) of the outputs of all trees.
Gradient Boosting	When high predictive accuracy is desired (e.g., predicting house prices).	It uses an ensemble of weak prediction models, typically decision trees. The model is built in a stage-wise fashion, and each new model helps to correct the errors made by the existing ensemble.
Bagging Classifier	Similar to Random Forest but for use with any base estimator (e.g., anomaly detection, classification tasks).	It creates an ensemble of base estimators, each trained on a different randomly sampled subset of the data. The final output is decided by majority voting (for classification) or averaging (for regression).
AdaBoost Classifier	Used in conjunction with many types of learning algorithms to improve performance (e.g., face detection).	It adjusts the weights of the training instances: instances that were misclassified by the previous model are given more weight. Each model in the ensemble focuses more

Model Training and Evaluation

East Kent Hospitals University NHS Foundation Trust

- How do we evaluate models?

Metric	Example Use Cases	Why is it Important to Consider this Metric?
Accuracy	General classification problems	Accuracy tells us about the overall performance of the model, giving us a basic idea of how well the model performs across all classes.
Precision **Precision measures how many of the predicted positive instances were actually correct.**	When false positives are costly (e.g., spam detection)	Precision is crucial when the cost of false positives is high. It tells us about the model's performance with respect to false positives.
Recall **Recall measures how well the model identifies the positive instances out of all the actual positive instances.**	When false negatives are costly (e.g., cancer diagnosis)	Recall is crucial when the cost of false negatives is high. It tells us about the model's performance with respect to false negatives.
F1 Score **The F1-score is the harmonic mean of precision and recall**	When an equal emphasis is given to both precision and recall (e.g., information retrieval)	F1 Score is the harmonic mean of precision and recall, and gives a balanced measure of the model's performance when both false positives and false negatives are equally important.
ROC-AUC	When we want to measure the model's ability to distinguish between classes	ROC-AUC gives an aggregate measure of performance across all possible classification thresholds, providing a measure of how well the model can distinguish between classes.



$$\text{Precision} = \frac{\text{True Positive}}{\text{Actual Results}} \quad \text{or} \quad \frac{\text{True Positive}}{\text{True Positive + False Positive}}$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{Predicted Results}} \quad \text{or} \quad \frac{\text{True Positive}}{\text{True Positive + False Negative}}$$

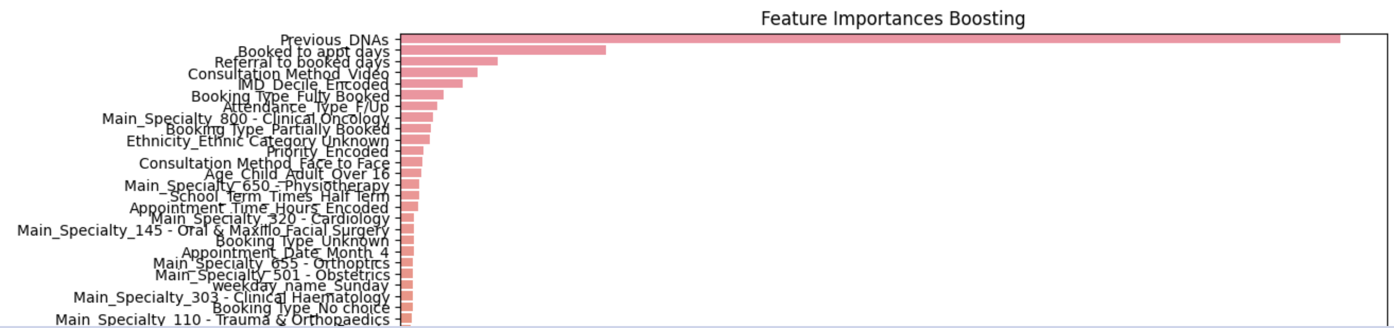
$$\text{Accuracy} = \frac{\text{True Positive + True Negative}}{\text{Total}}$$

$$F1\ score = 2 * \frac{Precision * Recall}{Precision + Recall}$$

NHS
East Kent
Hospitals University
NHS Foundation Trust

Model Training and Evaluation

- Important Features

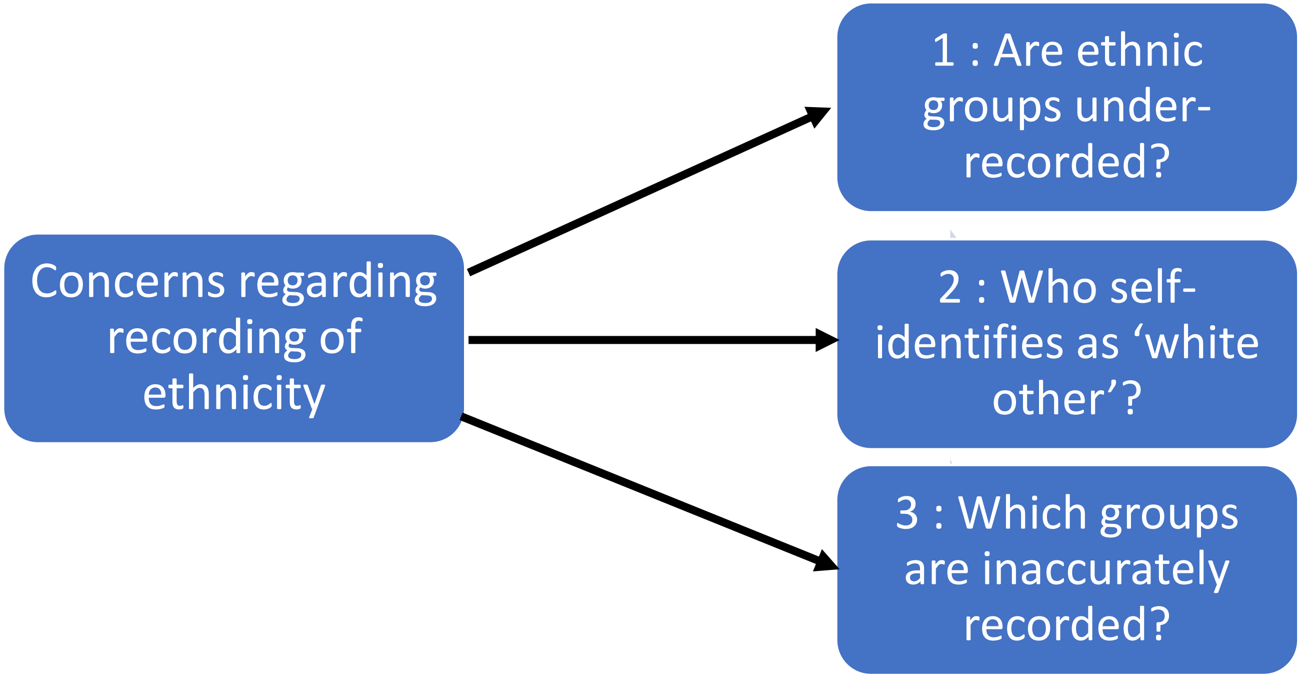
Model	Metrics
AdaBoost Classifier	

- Best Model selection

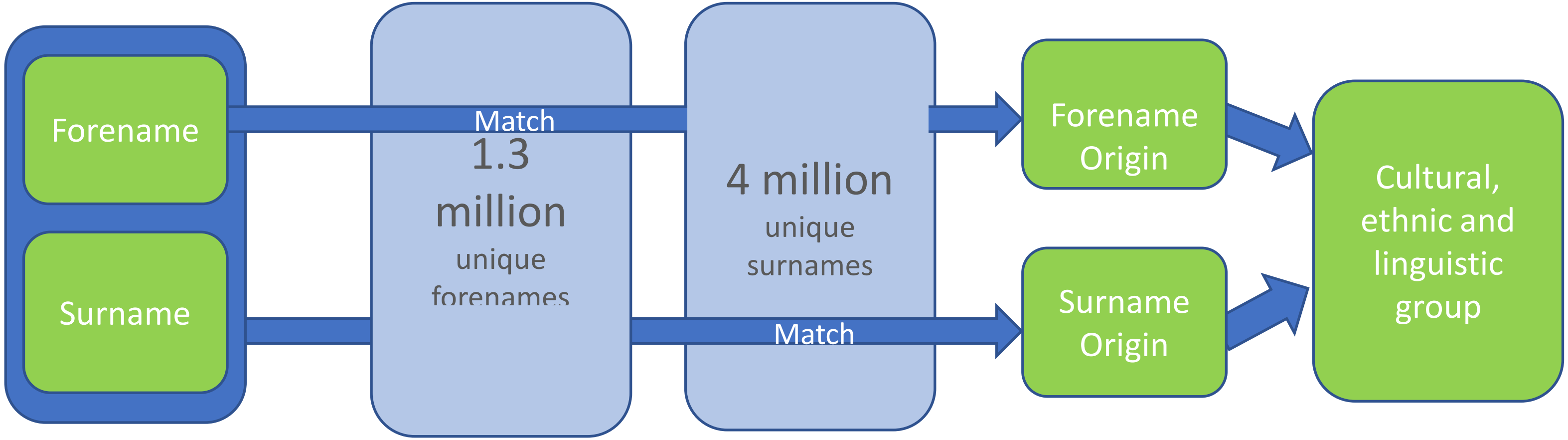
```python
# Accuracies Dictionary
accuracies = {
    "Logistic Regression": lr_accuracy,
    "Decision Tree": dt_accuracy,
    "Random Forest": rf_accuracy,
    "Gradient Boosting": gb_accuracy,
    "Bagging": bagging_accuracy,
    "AdaBoost": boosting_accuracy
}

# Choose the model with the highest accuracy
best_model_name = max(accuracies, key=accuracies.get)
best_model = models[best_model_name]
print(best_model)
```

```
RandomForestClassifier(random_state=42)
```

Concerns of East Kent Hospital Trust



Origins software



1 : Are ethnic groups under-recorded?

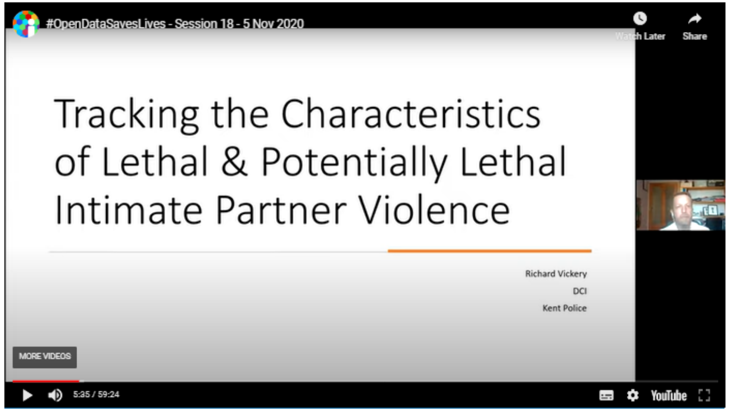
Origin of name	% patients ethnicity not known or patients declined to answer
Britain and Ireland	5.38
Baltic States	15.35
Nigeria	16.51
Romania or Moldova	18.02
Bangladesh	18.04

2 : Who self- identifies as ‘white other’?

Origins Sub Groups	patients	% describing themselves as British	% describing themselves as any other white background
EIA : POLISH	7,273	30.06	50.05
EJZ : CZECH OR SLOVAK	3,994	17.88	50.15
EKA : HUNGARIAN	1,850	34.22	34.92
ELZ : BALTIC STATES	1,824	16.50	57.95
EMF : ALBANIAN	966	33.75	41.61
EMZ : FORMERLY YUGOSLAV	1,404	39.60	33.97
ENA : BULGARIAN	1,024	17.68	53.42
EOA : ROMANIAN	4,279	10.56	56.65
EPZ : RUSSIAN OR UKRAINIAN	4,313	18.76	51.05

Session 18

Date: 5 November 2020, 12PM



Richard Vickery from **Kent Police** talks about the importance of sharing data and will also share the origins of a study in Kent focusing on domestic abuse.

Professor Chris Farmer from **Kent University** describes how we can link data from different organisations to create longitudinal studies of disease prevalence.

Open-source wizard **Chris Beeley** from **Nottinghamshire NHS** on why the government must share its code.

You can find notes from the session in the [Shared Google Doc]

[Richard Vickey, DCI, Kent Police]

Share

Resourcing our drive to reduce inequalities

Educating staff and creating resources

Increased Governance around inequalities monitoring

Use of AI

Professional accreditation

Health Inequalities Reporting

Created by Imogen Davies, last modified by Marc Farr on Mar 01, 2023

"Health inequalities can be defined as differences in health status or in the distribution of health determinants between different population groups" – World Health Organisation.

Inequality Variables

Main NHS focus: Ethnicity and Deprivation

Because ethnicity and deprivation were identified as being correlated with Covid-19 outcomes, these two variables are the main focus in NHS guidance currently (see 2022/23 Operational Planning guidance).

Resources on data quality/recording:

- NHSE/I Ethnicity and Deprivation Information Standard 2021
- Data Dictionary: Ethnic Category

Challenges:

- NHS Race & Health Observatory 2021 report: Data quality issues in hospital ethnicity data
- NHS ethnic categories are still based on 2001 Census rather than 2011/2021
  - 'Chinese' is under Other in 2001, rather than Asian 2011 onwards
  - 'Arab' doesn't exist yet in 2001
  - 'Gypsy or Irish Traveller' doesn't exist yet in 2001
- Deprivation uses IMD-2019, which categories patients according to the average deprivation in their local area, rather than their own personal level of deprivation.
  - May already be out of date.
  - Other measures e.g. Occupation or Educational Attainment Level are better for individual level of deprivation, but harder to get data.

Protected characteristics

There are 9 "protected characteristics" in UK law (Equality Act 2010). These are:

- Age
- Disability
- Gender reassignment
- Marriage and civil partnership
- Pregnancy and maternity
- Race
- Religion or belief
- Sex
- Sexual orientation

As public bodies, NHS organisations have a duty to avoid discrimination against these groups, and NHS England guidance mentions that inequalities for these characteristics should be monitored where relevant/appropriate.

Resources on data quality/recording:

- NHSE/LGBT Foundation guidance on monitoring sexual orientation and trans status 2021
  - Full specification here
- Data Dictionary: Disability
- ... more to add

Challenges:

- In NHS data, disability includes people who consider themselves to be disabled, as well as people diagnosed with a condition that means they might consider

Other "key groups"

This is a list of all the other groups mentioned in NHS England guidance here and new NHS Engalnd Health Inequalities reduction programme, Core20Plus5. Sometimes called "Inclusion Health Groups".

- Learning Disability
- Severe Mental Illness
- Frailty
- Homeless people
- Gypsy/Roma/Traveller
- People in prison or "in contact with the justice system"
- Immigrants / "vulnerable migrants"
- Sex workers
- Ethnic minority groups
- Coastal communities
- People with multi-morbidities
- People with drug and alcohol dependence
- Victims of modern slavery
- "Other socially excluded groups"

Idea is for local areas (ICS-level) to set their own priorities on which groups to focus on from this list.

Resources:

- Coding of homelessness
- Postcodes of prisons and immigration detention centres 2021-22
- Chief Medical Officer 2021 report on health challenges in Coastal Communities

NHS
East Kent
Hospitals University
NHS Foundation Trust

TERMS OF REFERENCE

INEQUALITIES AND UNWARRANTED VARIATION COMMITTEE

1. CONSTITUTION

1.1 This Committee is constituted by CEMG as the senior operational group responsible for addressing inequalities.

1.2 The EKHUFT We Care programme has a strategic initiative for 'Sustaining Access' and within that to do it equitably. This working group will put in place the analysis capability for the trust to be able to deliver on the commitment in the operational planning guidance and consider what actions should be carried out to address any inequalities.

1.3 These terms of reference set out the committee's response to the remit in terms of its constitution, governance, operation and activities.

2. PURPOSE

The Working Group shall:

2.1 Provide an analytical framework for analysing inequalities and variation in the

Artificial Intelligence & Robotic Process Automation

CogStack

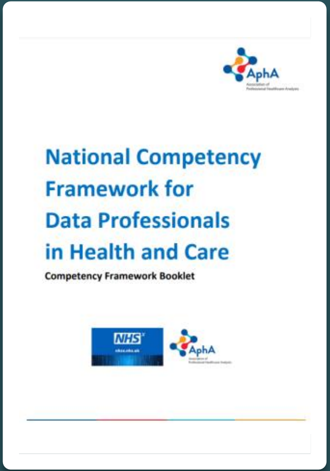
AUTOMATION ANYWHERE

AI CENTRE for Value Based Healthcare

Developing the Analysts of tomorrow

skills development network

Graduates into Health


AphA

National Competency Framework for Data Professionals in Health and Care

Competency Framework Booklet

FutureNHS

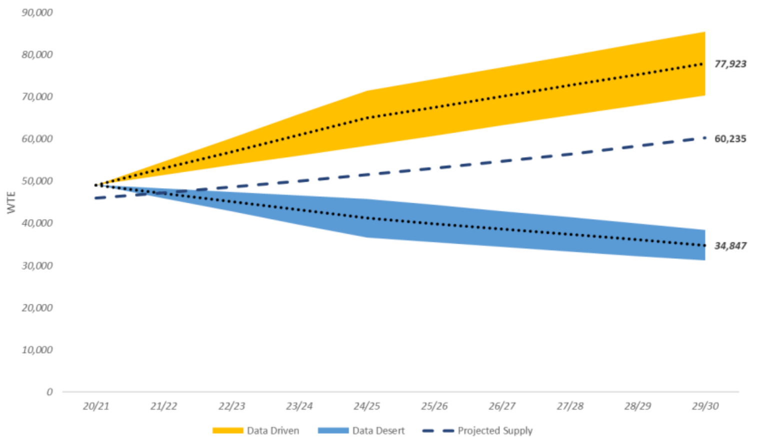
Microsoft Learning

DAVIES FURLONG CONSULTING

NHS
East Kent Hospitals University NHS Foundation Trust

Avoiding the data desert

Figure 1: Supply projection and demand forecasts for the NHS digital technology and health informatics workforce in a Data Driven Future and Data Desert Future – 2020/21 to 2029/30



Source: HEE Data Driven Healthcare in 2030

NHS
East Kent
Hospitals University
NHS Foundation Trust

Re-imagining Health Equity

Prof. Marc Farr
Chief Analytical Officer,

East Kent Hospitals NHSFT
Kent and Medway ICB

NHS
East Kent
Hospitals University
NHS Foundation Trust

INTELLIGENT HEALTH UK 2023

Breaking down the barriers between tech and healthcare