IMPROVING CROWDSOURCING OF HISTOPATHOLOGY ANALYSIS THROUGH TUTORIALS AND USER WEIGHTING

Rupesh Robinson-Vyas[1], Janette McQuillan[1], Jonathan Lawson[1], Dryden Williams[1], Andy Paterson[1], Sarah Christie[1], Josh Lee[1], Leslie Harris[1], Manuel Salto-Tellez[2], Paul Pharoah[3], Diana Eccles[4], Karwan Moutasim[4], Emily Shaw[1,4], Anne Kiltie[5], Will Howat[3], Gareth Thomas[4], Andy Hanby[6] and **Peter Smittenaar[1]**

[1] Cancer Research UK Citizen Science Team , [2]Queen's University Belfast, [3]CRUK Cambridge Institute, [4]University of Southampton , [5]CRUK/MRC Oxford Institute for Radiation Oncology, [6]Leeds University

INTRODUCTION

Given effective training, members of the public can perform high volume analysis of research data. This study investigates the efficacy of various teaching methods at training members of the public to accurately discriminate between cancer cells and non-cancer cells in images of immunostained tumour samples (TMA cores) using our web application, Trailblazer.

MATERIALS AND METHODS

Experiment design

Users were screened to exclude any with advanced science training.

They were then randomly allocated to one of four tutorials featuring annotated example images, interactive feedback, both or neither.

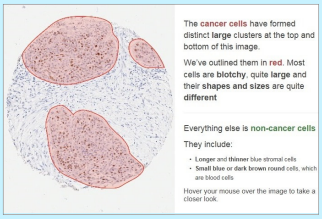
Users then conducted analysis on 10 sample images. Their performance was compared against consensus expert analysis to determine accuracy.

TRAILBLAZER EXPERIMENT DESIGN

	Annotated Images: Tutorial includes five images with cancer and non-cancer cells outlined and descriptions to guide identification — YES	NO
Feedback: Tutorial includes five practice images with interactive feedback and learning-dependent progress. — YES	n = 66	n = 65
NO	n = 66	n = 65

TUTORIAL COMPONENTS

Annotated examples



Interactive feedback

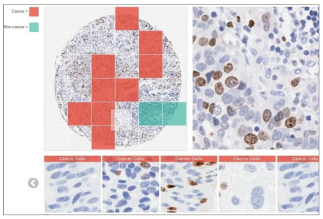
Trailblazer

Our web hosted platform, guided users through training and testing.

Users required no special software and accessed Trailblazer by clicking a link emailed to them.

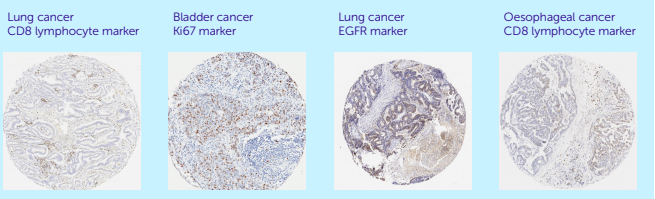
Each sample image was divided into 36 sections to provide more detailed results.

Analysis



Samples

Images of 0.6mm TMA cores, scanned at x40 magnification from the following tumour/marker type combinations:



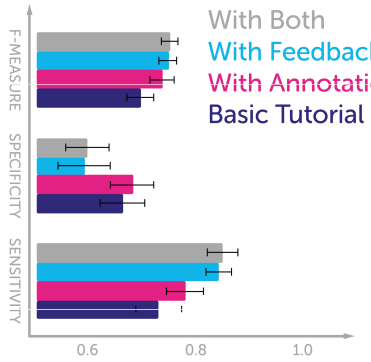
CONCLUSION AND NEXT STEPS

Members of the public can be trained to detect cancer cells in TMA cores. Providing feedback on performance improves accuracy.

Our next step is to investigate whether members of the public can accurately interrogate the nature of immunostains present within cancer cells in images containing a mixture of cell types and artefacts. If successful, this could provide a powerful analytical tool for high throughput image analysis in cancer research.

RESULTS

Improving user accuracy: CD8 lymphocyte stained lung cancer samples

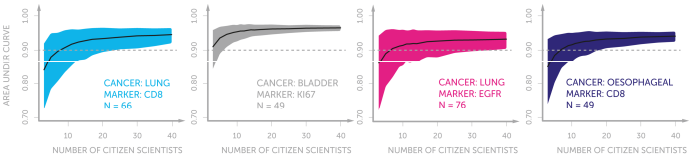


Annotated images and feedback both improved average accuracy (F-measure) at detecting cancer cells (p=0.002).

Provision of feedback had a more significant effect on accuracy than annotated images.

There is little interaction effect between the two methods.

User performance: Across four tumour/marker type combinations

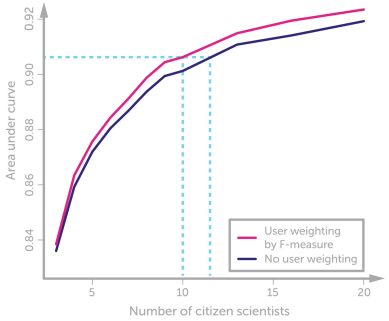


Aggregating the analysis of 10 or more participants trained with both feedback and annotated examples, achieved accurate detection of cancer cells (Area Under Curve >0.90) for all samples tested.

User-weighting to improve accuracy

Having 10 users after user weighting is equivalent in AUC to 11.5 users without weighting.

That means user weighting reduces required human input by 13%.



cruk.org/citizenscience

CANCER RESEARCH UK CITIZEN SCIENCE