Automated detection of Mycobacterial growth on 96-well plates for rapid and accurate Tuberculosis drug susceptibility testing.

Philip W. Fowler*[1,2], Ana Luíza Gibertoni Cruz[1], Sarah J. Hoosdally[1], Lisa Jarrett[3], Emanuele Borroni[4], Matteo Chiacchiaretta[4], Priti Rathod[3], Timothy M. Walker[1,2], Esther Robinson[3], Timothy E. A. Peto[1,5], Daniela Maria Cirillo[4], E. Grace Smith[3], and Derrick W. Crook[1,5]

[1]*Nuffield Department of Medicine, John Radcliffe Hospital, University of Oxford, Headley Way, Oxford, OX3 9DU, UK*

[2]*National Institute for Health Research (NIHR) Oxford Biomedical Research Centre, John Radcliffe Hospital, Headley Way, Oxford, OX3 9DU, UK*

[3]*National Mycobacterial Reference Service, National Infection Service, Public Health Laboratory Birmingham, Heartlands Hospital, Bordesley Green, Birmingham, B9 5SS, UK*

[4]*Emerging Bacterial Pathogens Unit, Division of Immunology, Transplantation and Infectious Diseases, IRCCS San Raffaele Scientific Institute, Milan, Italy*

[5]*NIHR Health Protection Research Unit in Healthcare Associated Infection and Antimicrobial Resistance at University of Oxford in partnership with Public Health England, Oxford, UK*

Abstract

Both liquid and solid culture-media based methods for drug susceptibility testing of *Mycobacterium tuberculosis* are slow, labour-intensive, expensive, and not fully standardised for all drugs. Microtitre plates have the potential to reduce cost and improve throughput. We have designed a 96-well microtitre plate containing serial

*To whom correspondence should be addressed: philip.fowler@ndm.ox.ac.uk

dilutions of 14 anti-tuberculosis drugs, and two control wells. Such plates have hitherto required reading by experienced operators; here we present some software, the Automated Mycobacterial Growth Detection Algorithm (`AMyGDA`), that identifies which wells in the plate contain bacterial growth, thereby elucidating the minimum inhibitory concentrations (MICs). We demonstrate that `AMyGDA` is consistent, and that the MICs measured are comparable to a reference method.

Tuberculosis (TB) kills more people globally than any other infectious disease (1, 2). In 2016 the World Health Organisation (WHO) estimated that only 22% of the 600,000 people who required treatment for multi-drug resistant tuberculosis (MDR-TB) were diagnosed and received appropriate therapy (1). In order to control the epidemic, priority should be given to the fast detection of MDR-TB cases and, via drug-susceptibility testing (DST), the identification of a proper and effective therapeutic regimen (3). One DST approach is to measure the minimum inhibitory concentrations (MICs) of a panel of antibiotics against a clinical isolate; the MIC is the lowest concentration that inhibits the growth of the pathogen. If any of the MICs are greater than their critical concentration, the strain is classified as being resistant to that antibiotic.

Existing liquid and solid media culture-based DST methods require significant infrastructure and highly-trained laboratory scientists. Although genotypic assays such as the Cepheid Xpert MTB/RIF (4) have been developed for *M. tuberculosis* (MTB) that overcome some of these challenges, culture-based DST remains essential to inform the use of both existing and future molecular tests: existing tests remain effective only as long as relatively few key mutations account for most drug-resistance whilst future tests will need to take into account the evolving background of genetic mutations that confer drug-resistance. A key way to monitor and map the ever-changing landscape of MTB genetic diversity is through direct measurement of the effect of different anti-MTB drugs on clinical samples taken from patients.

Microtitre plates are an attractive option for high-throughput culture-based DST, as they enable testing of a large number of strains in the presence of different anti-MTB drugs at a range of concentrations. One example is the Thermo Fisher Sensititre™ *M. tuberculosis* MIC Plate (MYCOTBI), which assays the minimum inhibitory concentration for 12 anti-MTB drugs. Following inoculation of a cultured isolate, a plate is incubated for two

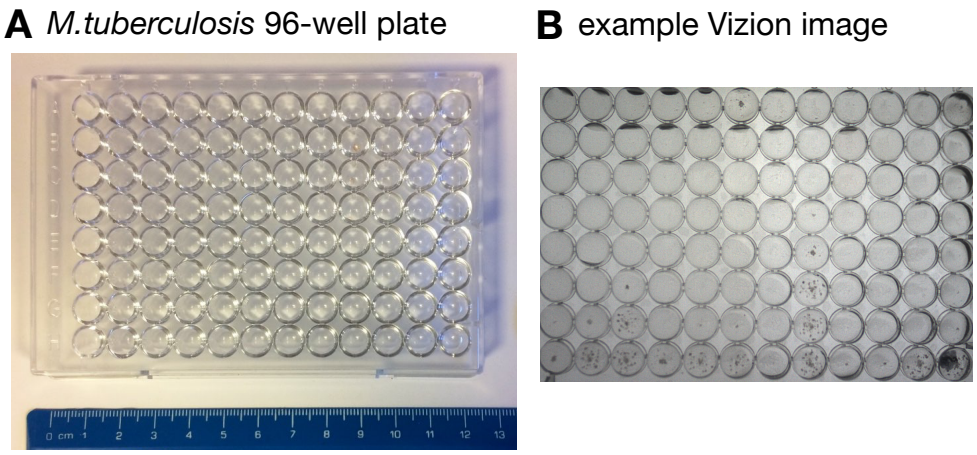

Figure 1: The 96-well plate used for *M. tuberculosis* (MTB) drug susceptibility testing. (**A**) The 96 well plate contains 14 different anti-MTB drugs. Each drug is present at 5-8 doubling dilutions and there are also two positive control wells which contain no drug. See Fig. S1 for the locations and concentrations of all 14 anti-MTB drugs. (**B**) Following inoculation with H37Rv and two weeks of incubation, an image is taken of each plate using a ThermoFisher Vizion™ Digital MIC viewing system. An example photograph is shown here. Specular growth can be seen in some wells. Raw images of all twenty plates can be found in Fig. S2

weeks, and the presence or absence of growth in each well, and hence the MIC, is then manually assessed by a trained laboratory scientist. The accompanying Sensititre™ Vizion™ Digital MIC viewing system is designed to help with this process.

The Comprehensive Resistance Prediction for Tuberculosis: an International Consortium (5) (CRyPTIC) is collecting over 30,000 samples of MTB world-wide over the next three years to identify the majority of genetic variation in MTB responsible for antibiotic resistance. CRyPTIC has developed a variant of the standard MYCOTBI plate, called the CRyPTIC1 plate, containing 14 different anti-MTB drugs, including additional second line drugs (linezolid and clofazimine) and two new compounds, bedaquiline and delamanid (Fig. 1A & S1). The drugs are present at a minimum of 5 and a maximum of 8 doubling dilutions. Each MTB sample collected will have its drug susceptibility profile determined using this plate. Any system based on visual assessment by a trained laboratory staff member is, however, subject to variability between operators. Automated reading, using specifically-designed computer software, offers the promise of greater consistency but may not be as accurate, given a human's superior ability to recognise and discriminate visual patterns.

In this paper, we describe the design and parametrisation of some software that detects the growth of MTB on 96-well plates, the Automated Mycobacterial Growth Detection Algorithm (`AMyGDA`). To test its reproducibility and accuracy, we apply it to twenty CRyPTIC1 plates inoculated with the H37Rv ATCC 27294 MTB reference strain (6).

MATERIALS AND METHODS

Study Design

Two TB Reference laboratories (Birmingham, UK and Milan, Italy) received one vial of MTB reference strain H37Rv. Each lab sub-cultured the isolate and inoculated ten CRYPTIC1 plates. After two weeks of incubation, the MICs of each plate were read by eye and then analysed by the `AMyGDA` software

Culturing

Each laboratory received one screw-capped plastic cryovial with approximately 1000 μl of mycobacterial suspension, and sub-cultured 200 μl in a mycobacterial growth indicator tube (MGIT; Becton Dickinson). A second subculture step was then performed: 200 μl of a well-mixed MGIT broth was inoculated on a solid Löwenstein-Jensen (LJ) medium, and incubated at 37 °C for 3-5 weeks. The MTB reference strain H37Rv, American Type Culture Collection 27294, obtained from the Biodefense and Emerging Infections Research Resources Repository (BEI Resources), USA was used throughout.

Plate inoculation

2 to 5 mg of growth was harvested from LJ media after 3 to 5 weeks incubation, and re-suspended in 5 mL of saline solution, homogenized, and the supernatant density adjusted to a McFarland standard 0.5 suspension ($\sim 1.5 \times 10^7$ CFU/mL). 100 μl of this suspension was inoculated into a MGIT and 100 μl dispensed from the MGIT into each of the plate's 96 wells. Each plate was covered with an adhesive seal and transferred to aerobic incubation at 35-37 °C

Plate reading by the Vizion™ Digital MIC viewing system

Fourteen days after inoculation, laboratory staff read the plate using a Thermo Fisher Vizion™ Digital MIC reading system. MICs were recorded in a database. Vizion™ images were then generated according to the manufacturer's guidance. Various combinations of Vizion™ lighting parameters – colour of background and level of illumination – were tested and, for the purposes of this study, the combination of 'white background' and 'level 7', respectively, was found to produce images with the highest contrast. Each image was manually cropped so it would fit on the Vizion™ Read window screen, with the result that final images had as similar dimensions as possible. Images were stored locally as lossless bitmap files (Fig. S2) by both laboratories before uploading to the database. In the Birmingham laboratory, the plate was then removed from the Vizion™ and a second laboratory scientist re-inserted the plate and took a second photograph, which was also uploaded to the database.

Software

AMyGDA is written in object-oriented Python. Each image is stored as a numpy array (7) and all image processing is done using the OpenCV2 Python API (8). To allow for straightforward metadata storage we use the datreant module (9). The algorithm assumes that each well is inoculated in the centre and mycobacteria will grow out radially from the centre, and that the images are 8-bit greyscale, allowing quantification of intensity of each pixel (range 0 to 255). Dark pixels (with low numerical values) were assumed to represent bacterial growth and light pixels to represent no growth. Growth was quantified according to the number of dark pixels per well.

The AMyGDA software can be freely downloaded from http://fowlerlab.org/software/amygda, subject to the end-user agreeing to an academic use software licence. The package contains instructions on how to install AMyGDA, including the installation of any pre-requisites, and also a short tutorial, which includes some sample images. AMyGDA is designed to be general, so that a different plate design is simple to implement and apply.

Filtering, image processing and growth detection

The raw images tend to be noisy, lacking in contrast and unevenly illuminated (Fig. 2A, S2 & S3A). To correct for these problems a mean shift filter (10) is applied (Fig. 2B, S3B & S4), followed by a contrast limited adaptive histogram equalization (CLAHE) filter (Fig. 2C, S3D & S5). The positions of the wells are then detected using a Hough transform optimised for circles as implemented in OpenCV; this is applied iteratively until only 96 circles are detected in the image. Next, the histogram of pixel intensities in a central square region of each well is calculated and the proportion of bacterial growth inferred. If above a specified threshold, the well is labelled with a coloured square that also defines the region analysed for growth. This process, including the choice of parameters, is described in detail in the Supplementary Information. Finally, in addition to saving the MICs to disc, each image is annotated with the name and concentration of the drug in each well, providing a final composite image (Fig. 2D, S12 & S13) that forms a complete, auditable record of the process.

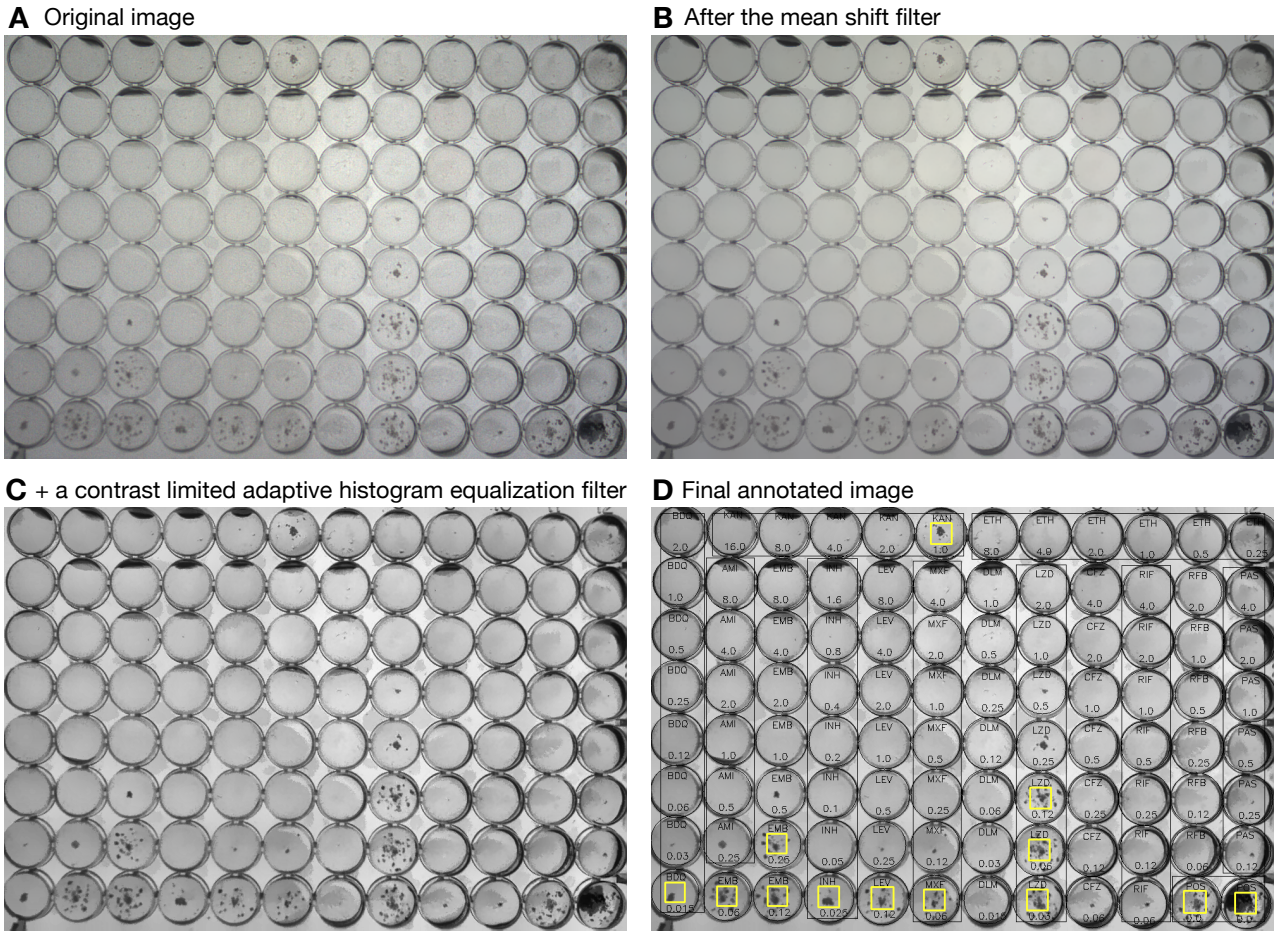

Figure 2: The AMyGDA software produces an annotated composite image with all detected growth labelled. (**A**) The original greyscale image is noisy and has poor contrast. To correct for this, (**B**) a mean shift filter and then (**C**) a contrast limited adaptive histogram equalization filter are applied. (**D**) The wells with detected bacterial growth are marked with a yellow square and each well is labelled with its drug and concentration and the estimated circumference is marked.

RESULTS

Sources of error

AMyGDA can wrongly report growth due to image artefacts, and miss existing growth where only small and/or faint patches of growth are apparent. As these sources of error are inversely coupled, a trade-off is inevitable. Artefacts which systematically affect a plate are more challenging; these include shadows (Fig. 3A) and remnants of the original inoculation ('sediment', Fig. 3B). Artefacts which occur more randomly result in nonsensical growth patterns – e.g. the bacteria appear to grow at high, but not low, concentrations of antibiotic – can be detected by the software and flagged. These include air bubbles (Fig. 3C), condensation (Fig. 3D), contamination (Fig. 3E) and possible failure of the integrity of a plate, e.g. evaporation of inoculum during incubation due to inadequate plate sealing (Fig. 3F). For more information please see the Supplementary Methods.

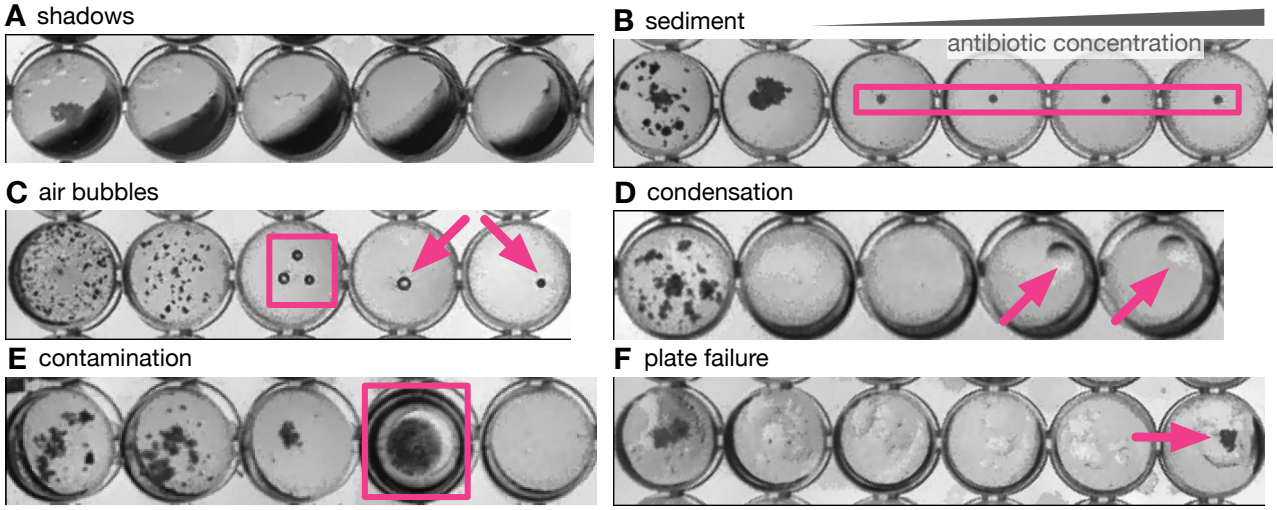

Figure 3: There are variety of artefacts that AMyGDA can mistake for growth, including (**A**) shadows, (**B**) sediment, (**C**) air bubbles, (**D**) condensation, (**E**) contamination and (**F**) possible failure of the plate integrity.

Reproducibility

AMyGDA produced identical results when applied ten times to the same image and is therefore consistent. When applied to two different images of the same plate the amount of detected growth in each well varied slightly from

one image to the other (Fig. 4A). This comparison only included a subset of wells in which H37Rv tends to grow well (both control wells and the lowest concentration wells for ethambutol, amikacin, isoniazid, levofloxacin, moxifloxacin and linezolid). Although a linear fit explains the data well, there are outliers and there is clearly some scatter as evidenced by a coefficient of determination of 0.93.

In 90% of cases (113/126) AMyGDA estimated an identical MIC (Fig. 4B & S14), including in six cases where a nonsensical growth pattern was detected in both images, namely where growth was detected in higher but not in lower concentration wells (Fig. 4C). Of the remaining 10%, the majority of these (7%) were only 1 doubling dilution different (Fig. 4C). These cases were a mixture of wells with little growth or artefacts. Taken together, this corresponds to an essential agreement of 97% (11). One case was 2 doubling dilutions different (Fig. 4D), due to condensation in this antibiotic lane being incorrectly classified as growth in two consecutive wells in one image, but not in the other. There were no cases where there was more than 2 doubling dilutions difference. The remaining 2% (3/126) cases were when the algorithm detected a nonsensical growth pattern in one image, but not in the other. Overall, AMyGDA satisfied the reproducibility requirement (95% of MICs being within 1 doubling dilution of the reference method) specified by the International Standards Organisation (11).

Validation

Prior to analysis by AMyGDA, all plates were read by experienced laboratory scientists using the Vizion™ Digital MIC instrument, allowing us to compare the MICs obtained by both methods (Fig. 5A). Both approaches rejected the same plate (1/20 plates, 5% drugs, Fig. 5B) due to lack of growth in one control well; in this case no MIC is returned for all 14 anti-MTB drugs on the plate. The AMyGDA software recorded nonsensical growth patterns for a further 11 (3.9%) drugs on other plates, thereby returning no MIC for 8.9% of drugs in total. The nonsensical growth patterns were mostly due to artefacts being incorrectly classified as bacterial growth (Fig. 5B). In all 11 cases the laboratory scientists were able to discern an MIC, demonstrating the superior ability of a human to identify and interpret artefacts. For 173/280 (62%) of drugs, both methods gave exactly the same result (Fig. 5C) and, overall, 242/280 (86%) measurements were within a single doubling dilution of each other (Fig. 5D). In 58/280 (21%) of cases AMyGDA reported an MIC one doubling dilution lower than the laboratory scientist. In contrast,

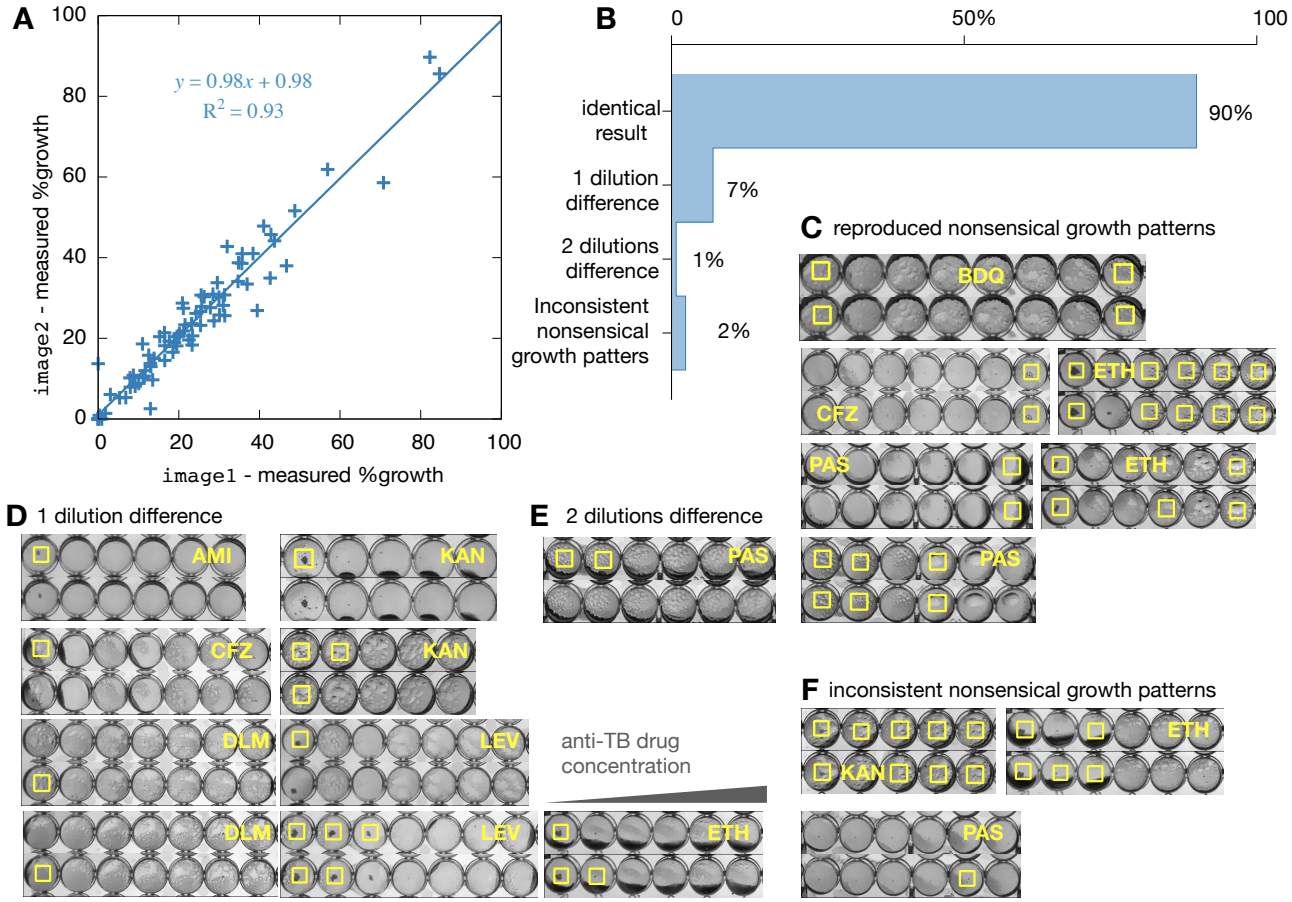

Figure 4: Applying the algorithm to the ten plates that were photographed twice (Fig. S14) measures the same MIC for 90% of wells. One of the ten replicates was excluded as the algorithm only detected growth in one of the two control wells. (**A**) The growth measured in a subset of wells on image1 and image2 is correlated. (**B**) The algorithm infers exactly the same MIC in 90% (113/126) of cases. This is comprised of 107 cases where the minimum inhibitory concentration is identical (examples not shown) and (**C**) six cases (5%) where there is a nonsensical growth pattern in both images. (**D**) There are nine cases (7%) where there is a single doubling dilution difference. (**E**) There is a single case where there are two doubling dilutions difference. (**F**) Finally, in three cases (2%) a nonsensical growth pattern is returned for one image, but not the other.

in only 11/280 (4%) of cases did the software predict an MIC one doubling dilution higher than the scientist. This disparity is due to the relatively conservative setting of the growth classification parameters to minimise the detection of artefacts. Excluding the plate rejected by both methods leads to an essential agreement (EA, MIC within ±1 doubling dilution) of 242/266 (90.9%) (11).

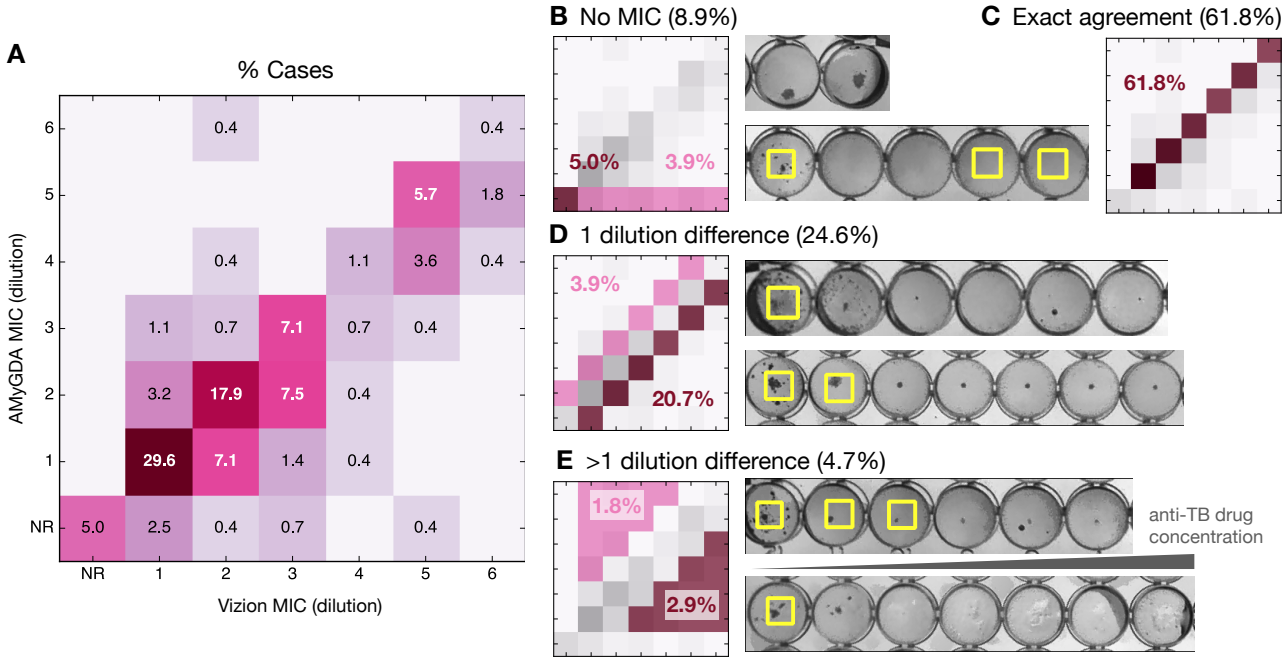

Figure 5: The results of the automated mycobacterial growth detection algorithm (AMyGDA) were validated by comparing to independent measurements made by laboratory scientists using the Vizion™ Digital MIC viewing instrument. (**A**) A heat map for all MIC doubling dilutions from all twenty test plates showing the concordance between the human-based measurement and the AMyGDA software. (**B**) Both measurements reject one plate (5%) which has no or unusual growth in one of the control wells. The AMyGDA software identifies nonsensical growth in a further 3.9% of cases. (**C**) Both methods infer the same MIC in 61.8% of cases with (**D**) a further 24.6% within ±1 doubling dilution. (**E**) In only 4.7% of cases do the methods disagree by more than a doubling dilution.

DISCUSSION

We have shown how `AMyGDA`, a Python-based code, can read images of a variant of the standard Thermo Fisher Sensititre™ Mycobacterial 96-well plates, each having been inoculated and incubated with the H37Rv *M. tuber-*

culosis strain. In a two-stage process AMyGDA first detects the locations of all 96 wells and then determines the presence (or absence) of bacterial growth, thereby measuring the minimum inhibitory concentrations for all 14 anti-MTB drugs on each plate. In this exploratory study, results were within the ranges set by the International Standards Organisation requirements for an MIC-based antibiotic susceptibility testing method (11). A larger and more comprehensive validation study is justified to evaluate accuracy and reproducibility of the AMyGDA software. AMyGDA can be downloaded from http://fowlerlab.org/software/amygda.

We envisage computer software like AMyGDA will supplement, rather than supplant, human-based plate reading. For example, *retrospective* analyses of plates would give useful information about the consistency and comparability of all staff members of a national clinical microbiology service and could also be used in training. More ambitiously, if every 96-well plate is *prospectively* read by software, plates could be flagged for further investigation if there was not sufficient agreement between the software and the laboratory scientist, reducing reading errors.

Since 96-well plates are potentially cheap to manufacture and distribute and as the software runs on a $40 Raspberry Pi computer, this combination could facilitate DST for second- and third-line anti-MTB drugs in high-incidence low- and middle-income countries where alternative culture-based DST approaches have traditionally been prohibitively expensive. An additional advantage is the provision of an audit trail, since all the metadata, including which wells were classified as containing growth, are embedded in the image (Fig. 2D, S12 & S13).

Conventionally culture-based methods are read after a fixed time period: here two weeks following inoculation. This implicitly assumes that all *M. tuberculosis* strains grow at a similar rate. We observed, however, significant variation in the amount of growth between plates even though these were all inoculated with the same strain, suggesting variability in inoculation has a potentially significant effect. This is likely to be accentuated further when different strains are compared. Software, like AMyGDA, could prospectively monitor the growth in the control wells during the incubation period, allowing a plate to only read once there is 'sufficient' growth ('read-when-ripe').

An additional use for automated plate-reading software could be for high-throughput phenotypic screening for MTB drug discovery. Unlike a human, the software is quantitative and so could, in theory detect small changes

in the rate of growth. Given enough samples, this could enable the use of microtitre plates to identify potential leads for novel antibiotic compounds by high-throughput phenotypic screening of a compound- or fragment-library which otherwise would be missed.

There are a number of limitations to our study. The Vizion™ was not designed to capture photographs that could be processed by software and hence it has uneven illumination and the digital camera is, by today's standards, of low resolution. An alternative, cheaper way of capturing images with more diffuse illumination and a higher resolution camera would help considerably. Whilst we have minimised the false detection of artefacts, they remain a problem, especially those that tend to systematically afflict whole regions of a plate. Improved experimental protocols may help avoid some of these (e.g. air bubbles) and a re-design of the plate layout would help with others (e.g. shadows). Using a circular (rather than square) measuring region would also help with the latter.

In conclusion, combining computer software, such as `AMyGDA`, with microtitre plate-based MIC measurement facilitates high-throughput culture-based drug susceptibility testing for tuberculosis.

ACKNOWLEDGEMENTS

We are grateful to members of the Oxford Doctoral Training Centre 2016/7 cohort for algorithmic suggestions.

FUNDING INFORMATION

The research was funded by the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre (BRC); the CRyPTIC consortium which is funded by a Wellcome Trust/Newton Fund-MRC Collaborative Award [200205/Z/15/Z] and the Bill & Melinda Gates Foundation Trust [OPP1133541]. T.E.A.P. and D.W.C. are NIHR Senior Investigators. T.M.W. is an NIHR Academic Clinical Lecturer. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

References

[1] **World Health Organisation**. 2017. Global Tuberculosis Report.
http://www.who.int/tb/publications/global_report/archive/en/

[2] **Fogel N.** 2015. Tuberculosis: A disease without boundaries. Tuberculosis **95**:527–531. http://dx.doi.org/10.1016/j.tube.2015.05.017.

[3] **World Health Organisation**. 2011. Towards universal access to diagnosis and treatment of multidrug-resistant and extensively drug-resistant tuberculosis by 2015. http://www.who.int/tb/publications/2011/mdr_report_2011/en/

[4] **Boehme CC, Nabeta P, Hillemann D, Nicol MP, Shenai S, Krapp F, Allen J, Tahirli R, Blakemore R, Rustomjee R, Milovic A, Jones M, O'Brien SM, Persing DH, Ruesch-Gerdes S, Gotuzzo E, Rodrigues C, Alland D, Perkins MD**. 2010. Rapid Molecular Detection of Tuberculosis and Rifampin Resistance. New Eng J Med. **363**:1005–1015. http://dx.doi.org/10.1056/NEJMoa0907847

[5] Comprehensive Resistance Prediction for Tuberculosis: an International Consortium (2017). http://www.crypticproject.org.

[6] **Kubica GP, Kim TH, Dunbar FP**. 1972. Designation of Strain H37Rv as the Neotype of *Mycobacterium tuberculosis*. Int J System Bacteriol **22**:99–106. http://dx.doi.org/10.1099/00207713-22-2-99

[7] **Van Der Walt S, Colbert SC, Varoquaux G**. 2011. The NumPy array: A structure for efficient numerical computation. Comp Sci Eng **13**:22–30. http://dx.doi.org/10.1109/MCSE.2011.37

[8] **Bradski G, Kaehler A**. 2008. Learning OpenCV. O'Reilly Media, Inc, first edition.

[9] **Dotson DL, Seyler SL, Linke M, Gowers RJ, Beckstein O**. 2016. datreant: persistent, Pythonic trees for heterogeneous data. In Proc 15th Python Sci Conf, edited by S Benthall, S Rostrup, 51–56.

[10] **Comaniciu D, Meer P**. 2002. Mean shift: A robust approach toward feature space analysis. IEEE Trans Pattern Analysis Mach Intel **24**:603–619. http://dx.doi.org/10.1109/34.1000236

[11] **International Organization for Standardization**. 2007. ISO 20776-2: Clinical laboratory testing and in vitro diagnostic test systems.