EBioMedicine 2 (2015) 1224–1234

Contents lists available at ScienceDirect

EBioMedicine

journal homepage: www.ebiomedicine.com

Research Paper

Do We Know Why We Make Errors in Morphological Diagnosis? An Analysis of Approach and Decision-Making in Haematological Morphology

Michelle Brereton [a], Barbara De La Salle [b], John Ardern [a], Keith Hyde [a,c], John Burthem [a,d,*]

[a] Central Manchester Foundation Trust, Oxford Road, Manchester M13 9WL, UK
[b] UK NEQAS(H), PO Box 14, Watford WD18 OFJ, UK
[c] School of Healthcare Sciences, Manchester Metropolitan University, John Dalton Building, M1 5GD, UK
[d] Institute of Cancer Sciences, 5th Floor St Marys Hospital, University of Manchester, M13 9WL, UK

ARTICLE INFO

Article history:
Received 6 June 2015
Received in revised form 8 July 2015
Accepted 14 July 2015
Available online 18 July 2015

Keywords:
Blood cell morphology
Diagnosis
Cell recognition
Heuristics
Errors

SUMMARY

Background: The laboratory interpretation of blood film morphology is frequently a rapid, accurate, and cost-effective final-stage of blood count analysis. However, the interpretation of findings often rests with a single individual, and errors can carry significant impact. Cell identification and classification skills are well supported by existing resources, but the contribution and importance of other skills are less well understood.
Methods: The UK external quality assurance group in haematology (UK NEQAS(H)) runs a Continued Professional Development scheme where large digital-images of abnormal blood smears are presented using a web-based virtual microscope. Each case is answered by more than 800 individuals. Morphological feature selection and prioritisation, as well as diagnosis and proposed action, are recorded. We analysed the responses of participants, aiming to identify successful strategies as well as sources of error.
Findings: The approach to assessment by participants depended on the affected cell type, case complexity or skills of the morphologist. For cases with few morphological abnormalities, we found that accurate cell identification and classification were the principle requirements for success. For more complex films however, feature recognition and prioritisation had primary importance. Additionally however, we found that participants employed a range of heuristic techniques to support their assessment, leading to associated bias and error.
Interpretation: A wide range of skills together allow successful morphological assessment and the complexity of this process is not always understood or recognised. Heuristic techniques are widely employed to support or reinforce primary observations and to simplify complex findings. These approaches are effective and are integral to assessment; however they may also be a source of bias or error. Improving outcomes and supporting diagnosis require the development of decision-support mechanisms that identify and support the benefits of heuristic strategies while identifying or avoiding associated biases.
Funding: The CPD scheme is funded by participant subscription.

1. Introduction

The United Kingdom National External Quality Assurance Scheme for Haematology (UK NEQAS(H)) provides a Continued Professional Development (CPD) scheme serving more than 2000 individuals within the UK and internationally, with around 1000 individuals completing each release. The scheme presents participants with large digital images of blood films within a software environment that mirrors elements of the microscope, and tests skills of feature identification and interpretation (Brereton et al., 2008). The outcome of interpreting morphology presented as digital slides is comparable with results using conventional microscopy of blood smears on glass slides (Burthem et al., 2005). The cases vary in complexity, and the level of skill or experience of participants differs, so the outcome of interpretation often varies markedly. We have assumed that errors of interpretation reflect lower levels of experience or knowledge, and that similar principles of interpretation and error apply across all cases. However, we have not previously tested whether these assumptions are correct. Drawing on evidence from other spheres of medicine, the present paper examines how our participants approach interpretation of blood film morphology and why that interpretation is sometimes incorrect.

Interpretation of blood films is a complex process: the first and central skill is the assignment of identities to the cells that are present (recognition and classification). If more than one cellular element is abnormal then the different features must be prioritised relative to each other (weighting). The goal in all cases involves an interpretation

* Corresponding author at: Clinical Senior Lecturer, Institute of Cancer Sciences, 5th Floor St Marys Hospital, University of Manchester, M13 9WL, UK.
E-mail address: john.burthem@cmft.nhs.uk (J. Burthem).

http://dx.doi.org/10.1016/j.ebiom.2015.07.020

of the findings (decision), but the nature of that decision may depend on level of skill and responsibility of the morphologist (expectation). In the laboratory setting, the time devoted to blood film examination is strongly influenced by the requirement to finish and move to next piece of work (completion). To help manage this complex decision process we consciously and unconsciously apply strategies that enable us to simplify and focus our analysis. The unconscious strategies are encompassed by the term "heuristics" (Shah and Oppenheimer, 2008). When they work well, heuristic approaches permit rapid and accurate interpretation: For example small children can readily and very rapidly identify different animals presented in pictures; by contrast computers struggle to reproduce this level of accuracy (Zhang et al., 2011). However, the "fast and frugal" heuristics that allow humans to outperform computers in many contexts can also be a source of bias (Marewski and Gigerenzer, 2012; Tversky and Kahneman, 1974). This bias frequently goes unrecognised by the individual, but may lead to mistaken conclusions and sometimes to serious error (Klein, 2005; Gunderman, 2009).

In the present study we have analysed the submissions of UKNEQAS(H) CPD participants assessing a range of representative cases, to examine the processes of decision making by individuals with varying levels of skill or experience. We have compared the submissions of morphologists arriving either at correct or incorrect diagnostic conclusions. Our analysis has revealed common patterns of approach to interpretation, but has also highlighted patterns of error shared by groups of participants. We suggest that our findings have relevance to the design of support mechanisms designed to improve the interpretation of haematological morphology.

2. Materials and Methods

2.1. Case Construction and Review

Cases were selected initially by members of the UK NEQAS(H) Morphology Scientific Advisory Group (Morphology SAG). Selection was based on blood smear quality, the range of morphological features, and the underlying diagnosis. Images were captured using a Zeiss Axio Imager M1 microscope and HRc camera (x63 Plan Apo Chromat 1.4 Oil immersion lens). At least 50 adjacent fields were manually focussed then formed into a single continuous image (photomerge function of Adobe Photoshop CS5). Post-processing included adjustment to image brightness and contrast, colour balance (Curves function) and sharpness (Unsharp mask) to ensure that reproduction matched the corresponding glass slide appearances, then images were uploaded to the viewing software (Digital SlideBox, Leica Biosystems).

2.2. Software System and Data Collection

Using the software virtual microscope as described (Burthem et al., 2005), participants were given brief clinical data, and viewed the image using magnification and navigation functions. Using a structured menu system, participants used a list of 74 features to select up to 5 morphological descriptors that they judged to best describe the blood film appearances then placed them in priority order. Participants were asked an additional single best-answer multiple choice question (most frequently "what would you do now?"); then had the option to suggest their preferred diagnosis using free-text entry.

2.3. Data Sorting and Analysis

If a single feature had a high diagnostic significance this was considered as a single element, otherwise observations reflecting the same pathological process were considered as a combined group (expressed as the mean number of selections and standard error of the mean (SEM)). A "priority score" was generated from the rank assigned: for single elements this was the rank assigned by the participant, for feature-groups this was the highest rank for any element of that group. Statistical evaluation employed GraphPad Prism software (v6.04): a comparison of feature selection or diagnosis employed contingency table analysis (Chi-square test: Fisher's exact test, two tailed analysis); priority scores for frequency of choice were compared between multiple groups using a non-parametric ANOVA test (Kruskal–Wallis test with multiple comparisons of means); for two sets of observations a two tailed Mann–Whitney test was employed. Significance is indicated in figures as follows: $^{*}p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$.

Funding: the CPD system was funded by participant subscription.

3. Results

3.1. Participants, Cases and Software

All participants were registered with the UK-NEQAS(H) Digital Morphology CPD scheme and principally comprised qualified UK Biomedical Scientists. Digital images presented in the virtual microscope software (Fig. 1a and b) were viewed by 715-1028 (mean 842) individuals with answers submitted using defined on-line criteria (Fig. 1c and d). Heat map analysis was employed in some studies; this analysis demonstrated that the users scanned the film area at low magnification before selecting specific areas for detailed examination at high magnification (Fig. 1e). This pattern is consistent with approaches to whole slide viewing previously shown by others (Raghunath et al., 2012). Five selected cases were analysed (Fig. 2a to e and Table 1).

3.2. Cases With a Predominant Single Morphological Feature

The morphological features present in cases 1 and 2 (Fig. 2a and b) affected a single cell type. Almost all participants correctly identified and prioritised the affected lineage (Fig. 3a and b). However, within that lineage the classification of the abnormal cells differed significantly between participants, and could be divided into distinct subgroups that were linked to the classification of the abnormal cell type (Table 1). For case 1, those answering the case correctly identified the abnormal cells as reactive lymphocytes, but other subgroups incorrectly reported the abnormal cells to be neoplastic, or reported the presence of both of neoplastic and reactive cells (Fig. 3c). Case 2 showed similar findings, with abnormal neutrophils being identified as the most significant feature by almost all participants. Those participants correctly interpreting the case classified the cells as having Pelger–Huet morphology, while incorrect groups selected either pseudo Pelger morphology (diagnosing myelodysplasia), or "left-shifted" morphology (assigning a reactive condition) (Fig. 3d).

In addition however, morphological features affecting other lineages also were consistently reported. Those correctly diagnosing the case made the fewest additional selections. Those participants diagnosing a reactive process more frequently (but incorrectly) reported reactive changes affecting other cell lineages. Where a neoplastic disorder was diagnosed, participants selected a higher number of morphological features, but did not identify supporting evidence from other cell lineages (Fig. 3c and d). For both cases, the preferred action selected by participants was clearly linked to their morphological interpretation, and the diagnosis of neoplasia was associated with a higher perceived importance for action (Fig. 3e and f).

3.3. Cases Combining Complex Morphological Features

Cases 3 and 4 had greater morphological complexity. Case 3 demonstrated a microangiopathic haemolytic anaemia (MAHA) together with reactive lymphocytes, reflecting an actual pathological diagnosis of thrombotic thrombocytopenic purpura (TTP) arising during acute human immunodeficiency virus (HIV) infection (Fig. 2c and Table 1). Consistent with this increased complexity, participants reported a