

The effect of clinical history on accuracy of electrocardiograph interpretation among doctors working in emergency departments

Monique F Cruz
BSc(Med), MB BS,
Emergency Registrar¹

James Edwards
MB BS, MMedEd, FACEM,
Emergency Physician¹

Michael M Dinh
MB BS, MPH, FACEM,
Emergency Physician¹

Elizabeth H Barnes
MStat,
Biostatistician²

¹Emergency Department,
Royal Prince Alfred
Hospital, Sydney, NSW.

²NHMRC Clinical
Trials Centre,
University of Sydney,
Sydney, NSW.

cruziin@gmail.com

MJA 2012; 197: 161–165
doi: 10.5694/mja11.11598

Accurate interpretation of electrocardiographs (ECGs) is essential for doctors to avoid diagnostic errors and delays in time-critical management. Diagnostic errors can result from knowledge deficits but also from cognitive biases and heuristics (mental shortcuts).^{1–4} Clinical reasoning consists of two distinct, although not exclusive, types of mental processes — analytical and non-analytical.¹ In emergency departments (EDs), doctors are often required to interpret results of diagnostic tests such as ECGs with limited or evolving clinical information, and adapt by using non-analytical clinical reasoning strategies such as pattern recognition and heuristics.⁴ Heuristics have been associated with enhanced diagnostic accuracy,⁵ but may also predispose to diagnostic error caused by bias.^{4,5} The major cognitive biases in medicine are anchoring bias (tendency to fixate on first impressions), confirmation bias (seeking features to confirm rather than refute a diagnosis), and premature closure (accepting a diagnosis before it has been fully verified).⁴

Access to clinical history can both enhance diagnostic accuracy and increase diagnostic errors.^{6–11} A systematic review of the interpretation of radiology and pathology test results has shown that adding clinical history is associated with improved diagnostic accuracy.⁶ Some studies of ECG interpretation have shown that providing a clinical history or vignette has no significant effect on diagnostic accuracy,^{10,11} but others have shown that history suggestive of the correct diagnosis increases diagnostic accuracy and misleading history decreases diagnostic accuracy.^{7–9}

We aimed to determine whether bias in the clinical history affects the accuracy of ECG interpretation by ED doctors.

Abstract

Objective: To investigate whether bias in clinical history affects accuracy of electrocardiograph (ECG) interpretation among doctors working in emergency departments.

Design and setting: Observational study conducted at four teaching hospitals in Sydney from May to September 2011.

Participants: Emergency registrars and physicians (senior doctors), and interns, resident medical officers and senior resident medical officers (junior doctors).

Intervention: Participants interpreted 30 ECGs representing 10 diagnoses. ECGs were provided with positively biased history (suggestive of the correct diagnosis), negatively biased history (suggestive of an alternative diagnosis) or no history.

Main outcome measures: Accuracy of ECG interpretation, measured as a score out of 10 (for each category of clinical history) and as a percentage of correctly interpreted ECGs.

Results: Of 307 doctors who were sent a recruitment email for the study, 132 participated (43%). The overall mean accuracy of ECG interpretation was 52% (95% CI, 50%–53%). For junior doctors, mean accuracy was 42% (95% CI, 40%–44%); for senior doctors, it was 65% (95% CI, 62%–67%). In adjusted models, the mean predicted score for senior doctors provided with no history was 6.25 (95% CI, 5.90–6.62) with junior doctors obtaining mean scores 34% lower than senior doctors (95% CI, 29%–40%; $P < 0.001$). Compared with no history, positively biased history was associated with 42% higher mean scores (95% CI, 35%–49%; $P < 0.001$) and negatively biased history was associated with 34% lower mean scores (95% CI, 29%–39%; $P < 0.001$).

Conclusion: Bias in clinical history significantly influenced the accuracy of ECG interpretation. Strategies that reduce the detrimental impact of cognitive bias and improved ECG training for doctors are recommended.

Methods

We conducted an observational study of ECG interpretation by doctors working in EDs at four teaching hospitals in Sydney (Royal Prince Alfred Hospital, Liverpool Hospital, Concord Hospital, and Canterbury Hospital) during routine hospital-based teaching sessions from May to September 2011. Ethics approval was obtained from the ethics committees at all study sites.

Participants

A convenience sample of emergency physicians, emergency registrars, senior resident medical officers, resident medical officers and interns was invited to participate. A recruitment email was sent to all ED doctors before commencement of the teaching

sessions, and participation was voluntary. Participants provided written consent and were blinded to the study objective.

Demographic data collected included postgraduate year (PGY) and current position. Emergency physicians and registrars were classified as senior doctors, and interns (PGY 1), resident medical officers (PGY 2) and senior resident medical officers (PGY 3) were classified as junior doctors. Junior doctors had commenced, completed, or were due to undertake a 10-week emergency medicine rotation.

Intervention

ECGs were obtained from hospital patients, textbooks and internet-based sources. They were reviewed without clinical history by a cardiologist and a senior emergency physician,

1 Clinical histories provided for each electrocardiograph (ECG) diagnosis

Diagnosis	Positively biased history	Negatively biased history
Left bundle branch block	89-year-old man with shortness of breath and previous myocardial infarction	28-year-old man with anterior pleuritic chest pain and recent viral illness (suggestive of an alternative diagnosis of pericarditis)
Ventricular tachycardia	78-year-old man with decreased level of consciousness, pale and diaphoretic	22-year-old man with recurrent episodes of palpitations (suggestive of an alternative diagnosis of supraventricular tachycardia with aberrancy)
Supraventricular tachycardia	37-year-old woman with a history of recurrent palpitations, on verapamil	24-year-old man with a fever and hypotension (suggestive of an alternative diagnosis of sinus tachycardia)
ST elevation myocardial infarction (STEMI)	68-year-old woman with hypertension and diabetes with central heavy chest discomfort	26-year-old with pleuritic chest pain that is relieved on sitting forward (suggestive of an alternative diagnosis of pericarditis)
Brugada pattern	26-year-old man presents with syncope. Brother died suddenly aged 29	78-year-old man with crushing retrosternal chest pain (suggestive of an alternative diagnosis of myocardial infarction)
Right bundle branch block	38-year-old woman with pleuritic chest pain and left calf swelling	38-year-old Asian man with syncope (suggestive of an alternative diagnosis of Brugada pattern)
Complete heart block	74-year-old with dizziness and pre-syncope	25-year-old routine ECG before elective knee arthroscopy (suggestive of an alternative diagnosis of a normal ECG)
Wolff–Parkinson–White pattern	16-year-old boy presents after a syncopal event, awaiting ablative surgery	55-year-old homeless man found lying in gutter in winter (suggestive of an alternative diagnosis of Osborn waves associated with hypothermia)
Long QT interval	25-year-old woman presents following a syncopal episode; she is on multiple antipsychotic agents	50-year-old with abdominal pain, routine ECG while in department (suggestive of an alternative diagnosis of a normal ECG)
Pericarditis	25-year-old with sharp central chest pain and a recent viral illness	83-year-old man with crushing retrosternal chest pain, diaphoretic (suggestive of an alternative diagnosis of STEMI)

and were included in the study only if these doctors agreed on the diagnosis. Thirty ECGs representing 10 clinically important diagnoses were included (three ECGs per diagnosis). These diagnoses are shown in Box 1.

The 30 ECGs were randomly divided into three packs of 10, such that there was one example of each diagnosis per pack. Each ECG was preceded by positively biased, negatively biased, or no clinical history. Positively biased history was suggestive of the correct diagnosis and negatively biased history was suggestive of an alternative diagnosis (Box 1). A random number generator was used to order the ECGs within each pack and assign a different category of history to each ECG.

Participants received all 30 ECGs and were therefore exposed to three examples of each diagnosis in a random sequence of positively biased, negatively biased and no clinical history. They interpreted the ECGs under supervised test conditions and were given 1 hour to write the most likely “ECG diagnosis” for all 30 ECGs. They were instructed to provide a single diagnosis based on the ECG without descriptive interpretation, and completed and submitted each pack of 10 ECGs before progressing to the next pack. A post-test education session on ECGs (including answers) was conducted.

The ECG answers were marked independently by two emergency

physicians who were provided with a marking scheme and were blinded to the ECGs and study objective. Answers that included the correct diagnosis or terms listed in the marking scheme were assigned a score of one. Answers that listed more than one diagnosis, were simply descriptive or failed to provide a diagnosis were assigned a score of zero. No half marks were awarded. Inter-rater disagreements were adjudicated by two of us (JE and MFC).

Primary outcome

The primary outcome was the accuracy of ECG interpretation, measured as a score out of 10 and as the percentage of ECGs that were correctly interpreted. Scores out of 10 reflected the number of ECGs that were correctly interpreted by each participant for the 10 ECGs in each category of clinical history.

Statistical methods and analysis

This was an exploratory study and an a priori sample size calculation was not undertaken. An estimate of the required sample size was obtained from previous studies that investigated similar outcomes.^{7–12} A difference in mean scores of 1 out of 10 (ie, 10% difference) between clinical history categories was considered clinically important. Assuming a correlation coefficient of 0.50 between repeated observations on the same participant, a total group size of 129

would provide 80% power at 0.025 alpha level (significance level adjusted to account for multiple pairwise comparisons) to detect a difference in mean scores of 0.55.

For the primary outcome, standard deviations and interquartile ranges for scores and confidence intervals for percentages were calculated according to standard formulas but did not take into account the repeated observations on each participant, so these were used for descriptive purposes only. To account for repeated measures and distribution of results, a generalised estimating equation (GEE) model was used. Univariable and multivariable estimated effect sizes of clinical history and participant seniority were then obtained. An interaction term was included in the model to assess whether the effect of clinical history differed according to participant seniority. Estimated effect sizes from the GEE model were expressed as a ratio of scores between the group of interest and a reference group (intercept) with 95% confidence intervals and *P* values. For example, an effect size of 1.5 would indicate that the predicted mean score in one group was 50% higher than the predicted mean score in the reference group, and an effect size of 1.0 would indicate no effect. Statistical analyses were performed in PASW Statistics 18 for Windows (SPSS Inc, Chicago, Ill, USA).

Results

Of the 307 doctors who were sent the recruitment email, 132 participated in the study (43%) — 74 junior doctors (56%) and 58 senior doctors (44%).

The mean accuracy of ECG interpretation for all participants was 52% (95% CI, 50%–53%). For junior doctors, mean accuracy was 42% (95% CI, 40%–44%); for senior doctors, it was 65% (95% CI, 62%–67%). Participants were most accurate for the ECG diagnoses of ST elevation myocardial infarction and ventricular tachycardia and least accurate for Brugada pattern and long QT interval. The percentages of ECGs that were correctly interpreted for each diagnosis are shown in Box 2, and these data are further divided by clinical history in Box 3.

The percentages of ECGs that were correctly interpreted in each of the three ECG packs were similar: 53% (95% CI, 50%–55%) for pack 1, 53% (95% CI, 50%–55%) for pack 2 and 50% (95% CI, 47%–53%) for pack 3.

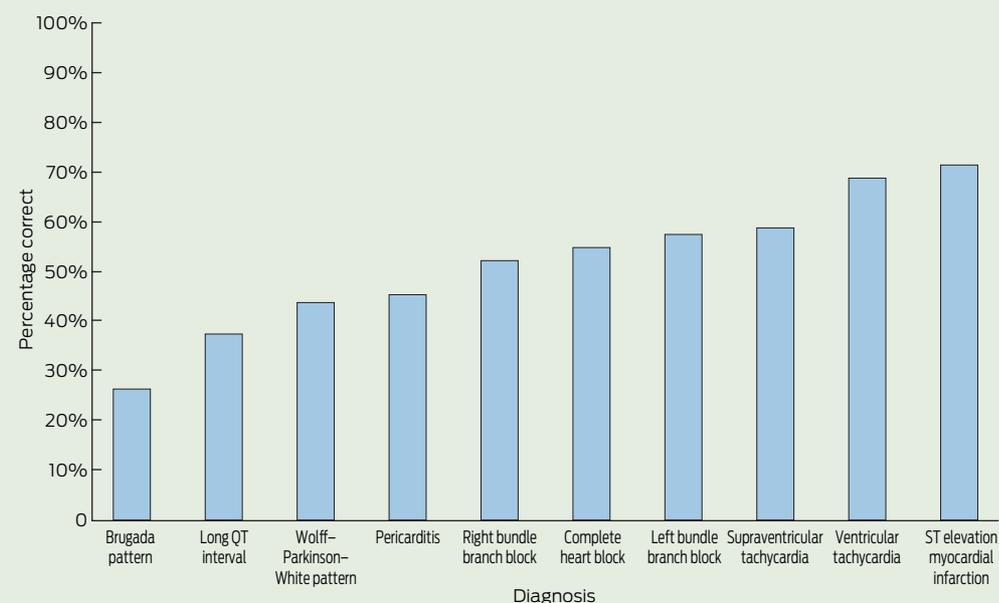
There were 97 marking disagreements out of the 3960 ECG answers marked, giving an agreement between markers of 98%.

Accuracy of ECG interpretation

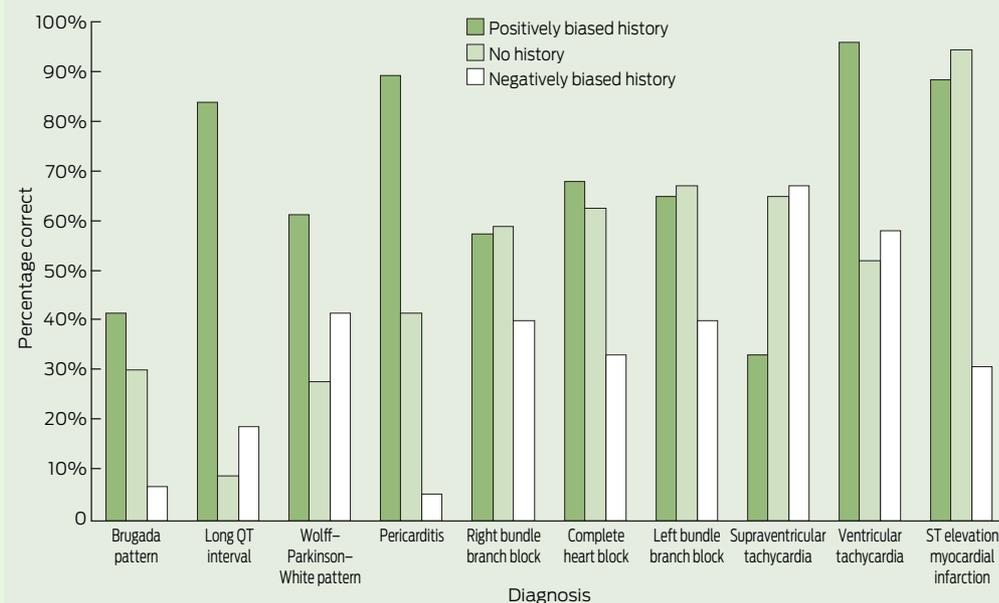
In multivariable analysis, there was no evidence of significant effect modification between clinical history and participant seniority (interaction term $P=0.21$). As a result, the interaction term was removed and only the main effects of history and seniority were included in the final model. There was a significant association between the category of history provided and the accuracy of ECG interpretation. Both mean and median test scores were significantly higher when the ECG was presented with positively biased history compared with no history, and lower when the ECG was presented with negatively biased history (Box 4).

In the final multivariable GEE model, the predicted mean score for senior doctors provided with no clinical history was 6.25 (95% CI, 5.90–6.62). Positively biased history was associated with a 42% (95% CI, 35%–49%) increase in mean test scores compared with no history. Negatively biased history was associated with a 34% (95% CI, 29%–39%) decrease in

2 Percentages of electrocardiographs (ECGs) that were correctly interpreted, by ECG diagnosis



3 Percentages of electrocardiographs (ECGs) that were correctly interpreted, by ECG diagnosis and clinical history



mean test scores compared with no history. Junior doctors' mean test scores were 34% (95% CI, 29%–40%) lower than those of senior doctors (Box 5). These effects are adjusted for other predictors in the model.

Discussion

Our study shows that introduction of bias into clinical history significantly affects diagnostic accuracy of ECG

interpretation. On one hand, doctors were more accurate when provided with history suggestive of the correct ECG diagnosis. The history may have directed participants to search for specific ECG findings or alerted them to possible diagnoses that they might not have otherwise considered. Clinicians may have also combined history with ECG findings to make their ECG diagnoses, rather than interpreting ECGs independently.¹³ This suggests

4 Observed mean and median test scores (out of 10), by clinical history and participant seniority			
	Positively biased history	No history	Negatively biased history
Senior doctors			
Median (interquartile range)	9.0 (8.0–10.0)	6.0 (5.0–8.0)	4.0 (2.8–6.0)
Mean (SD)	8.7 (1.3)	6.4 (1.6)	4.2 (2.1)
Junior doctors			
Median (interquartile range)	6.0 (5.0–7.0)	4.0 (3.0–5.0)	2.0 (1.8–4.0)
Mean (SD)	5.9 (1.7)	4.0 (1.7)	2.6 (1.5)
All doctors			
Median (interquartile range)	7.5 (5.0–9.0)	5.0 (4.0–6.8)	3.0 (2.0–5.0)
Mean (SD)	7.1 (2.1)	5.0 (2.0)	3.3 (2.0)

that ECG interpretation may not be independent of the pre-test probability, contrary to the traditional Bayesian approach to clinical reasoning.

On the other hand, clinical history suggestive of an alternative diagnosis had a detrimental effect on diagnostic accuracy. Cognitive processes such as anchoring bias, confirmation bias and premature closure may have contributed to participants ignoring findings consistent with the correct ECG diagnoses. This is consistent with results of research which show that diagnostic suggestion can lead to interpretation errors, particularly in the context of atypical or inconsistent clinical histories.^{14–16} Clinicians therefore need to ensure that any history obtained before the interpretation of ECGs is reliable and accurate. Furthermore, strategies are required so that an incorrect or atypical history does not adversely affect the interpretation of diagnostic test results.

Several such strategies have been proposed to reduce the risk of diagnostic error.^{1–4,9,13–18} One approach is to interpret test results without clinical history (to facilitate unbiased reading) and then reinterpret with the history.¹⁴ Another approach is to seek

more clinical information when ECG findings are inconsistent with history. A “combined reasoning approach” — the deliberate use of both analytical and non-analytical (intuitive) methods — has been associated with increased diagnostic accuracy for ECG interpretation.¹⁷ Another strategy is for clinicians to develop an understanding of basic error theory and skills in meta-cognition (ie, think about their thinking¹⁸), and use specific strategies to counter cognitive pitfalls.⁴ Despite multiple strategies being proposed, there is little evidence to support one approach over another.¹⁸

The strengths of this study include the blinded assessment of the ECG responses to minimise misclassification bias, the good inter-rater agreement, and the randomisation of ECGs within each pack to minimise the effect of learning bias. In contrast to previous similar studies,^{7–11} this was a multicentre study with a specific focus on the effect of clinical history and cognitive bias in the ED setting.

However, there were a number of limitations. There were not enough participants to enable subgroup analyses of each clinical designation. We did not control for other potential confounders such as hospital type or prior ECG training. Positively and negatively biased clinical histories provided were formulated by the authors and may have caused a bias in favour of the observed differences. Although the ECGs in our study were included only if both the emergency physician and cardiologist agreed on the diagnosis, significant inter-observer variability in ECG interpretation often exists even among cardiologists (who are frequently the “gold

standard” reference in similar studies).¹⁹ The accuracy of ECG interpretation observed in this study was probably lower than in clinical practice as a result of the strict marking scheme, designed to avoid ambiguity and marker bias. Some of the participants whose answers were marked as incorrect might have identified features in the ECGs that would have led to further enquiry in clinical practice. The clinical significance of these interpretation errors is unknown.

The decreased accuracy of ECG interpretation associated with negatively biased clinical history was consistent across senior and junior levels. In contrast, another study has shown that expertise can mitigate the detrimental effect of bias — providing a misleading history decreased diagnostic accuracy of ECG interpretation by 5% for cardiologists but up to 25% for residents.⁷ In our study, no such protective effect of expertise was seen, possibly because most of the senior doctors were emergency registrars and few were emergency physicians.

Overall accuracy of ECG interpretation in this study was 52%, which is comparable to findings of previous studies — accuracies of 36%–96% for emergency doctors.¹⁹ The performance of senior doctors was comparable to previously reported accuracies of 53%–96% for cardiologists.^{7,19} Accuracies for the individual ECG diagnoses in our study were similar to those reported in a recent Australian study.¹² The low diagnostic accuracy for some ECG patterns suggests that knowledge deficits may have been just as important as cognitive bias in contributing to diagnostic error. This is supported by the finding that the effect of seniority was similar to that of biased clinical history on the accuracy of ECG interpretation. Junior doctors may have been unable to identify ECG patterns of Brugada pattern and long QT interval due to a lack of prior exposure to these disorders. Knowledge gaps are more amenable to educational interventions and highlight that ongoing training in ECG interpretation is an important component of postgraduate education programs.

In our study, biased clinical history and participant seniority had significant effects on the accuracy of ECG inter-

5 Estimated effects of clinical history and participant seniority on mean test scores, from a generalised estimating equation model*

Predictor	Estimated effect (95% CI)	P
Positively biased history (relative to no history)	1.42 (1.35–1.49) [†]	< 0.001
Negatively biased history (relative to no history)	0.66 (0.61–0.71) [†]	< 0.001
Junior doctors (relative to senior doctors)	0.66 (0.60–0.71) [‡]	< 0.001

* Based on the generalised estimating equation model, the predicted mean score for the reference group (intercept) (senior doctors, no history) was 6.25 (95% CI, 5.90–6.62). [†] Adjusted for seniority. [‡] Adjusted for history category. ◆

pretation. These findings may be helpful in developing strategies to improve the accuracy of ECG interpretation by doctors working in EDs and other acute care settings. We recommend improved postgraduate training and strategies that limit the detrimental impact of cognitive bias.

Acknowledgements: We thank Raj Puranik (cardiologist) and Tim Royle (senior emergency physician) for reviewing the ECGs included in the study, and Samantha Bendall and Rosslyn Hing (emergency physicians) for marking the ECG answers.

Competing interests: Michael Dinh was awarded a grant by the New South Wales Government Health Education and Training Institute to develop and implement an online ECG training and knowledge-sharing platform.

Received 15 Dec 2011, accepted 2 Jul 2012.

- 1 Norman GR, Eva KW. Diagnostic error and clinical reasoning. *Med Educ* 2010; 44: 94-100.
- 2 Berner ES, Graber ML. Overconfidence as a cause of diagnostic error in medicine. *Am J Med* 2008; 121 (5 Suppl): S2-S23.
- 3 Graber ML. Educational strategies to reduce diagnostic error: can you teach this stuff? *Adv Health Sci Educ Theory Pract* 2009; 14 Suppl 1: 63-69.
- 4 Croskerry P. The importance of cognitive errors in diagnosis and strategies to minimize them. *Acad Med* 2003; 78: 775-780.
- 5 Sibbald M, Cavalcanti RB. The biasing effect of clinical history on physical examination diagnostic accuracy. *Med Educ* 2011; 45: 827-834.
- 6 Loy CT, Irwig L. Accuracy of diagnostic tests read with and without clinical information: a systematic review. *JAMA* 2004; 292: 1602-1609.
- 7 Hatala R, Norman GR, Brooks LR. Impact of a clinical scenario on accuracy of electrocardiogram interpretation. *J Gen Intern Med* 1999; 14: 126-129.
- 8 Hatala RA, Norman GR, Brooks LR. The effect of clinical history on physician's ECG interpretation skills. *Acad Med* 1996; 71 (10 Suppl): S68-S70.
- 9 LeBlanc VR, Dore K, Norman GR, Brooks LR. Limiting the playing field: does restricting the number of possible diagnoses reduce errors due to diagnosis-specific feature identification? *Med Educ* 2004; 38: 17-24.
- 10 Grum CM, Gruppen LD, Wolliscroft JO. The influence of vignettes on EKG interpretation by third-year students. *Acad Med* 1993; 68 (10 Suppl): S61-S63.
- 11 Dunn PM, Levinson W. The lack of effect of clinical information on electrocardiographic diagnosis of acute myocardial infarction. *Arch Intern Med* 1990; 150: 1917-1919.
- 12 Hoyle RJ, Walker KJ, Thomson G, Bailey M. Accuracy of electrocardiogram interpretation improves with emergency medicine training. *Emerg Med Australas* 2007; 19: 143-150.
- 13 Irwig L, Macaskill P, Walter SD, Housami N. New methods give better estimates of changes in diagnostic accuracy when prior information is provided. *J Clin Epidemiol* 2006; 59: 299-307.
- 14 Griscorn NT. A suggestion: look at the images first, before you read the history. *Radiology* 2002; 223: 9-10.
- 15 Leblanc VR, Norman GR, Brooks LR. Effect of a diagnostic suggestion on diagnostic accuracy and identification of clinical features. *Acad Med* 2001; 76 (10 Suppl): S18-S2.
- 16 Leblanc VR, Brooks LR, Norman GR. Believing is seeing: the influence of a diagnostic hypothesis on the interpretation of clinical features. *Acad Med* 2002; 77 (10 Suppl): S67-S69.
- 17 Eva KW, Hatala RM, Leblanc VR, Brooks LR. Teaching from the clinical reasoning literature: combined reasoning strategies help novice diagnosticians overcome misleading information. *Med Educ* 2007; 41: 1152-1158.
- 18 Scott IA. Errors in clinical reasoning: causes and remedial strategies. *BMJ* 2009; 338: b1860.
- 19 Salerno SM, Alguire PC, Waxman HS. Competency in interpretation of 12-lead electrocardiograms: a summary and appraisal of published evidence. *Ann Intern Med* 2003; 138: 751-760. □

Indigenous health: we're always interested

As part of the *MJA*'s redesign, we are seeking to keep the health of Aboriginal and Torres Strait Islander people on the agenda all year round.

We're looking to publish excellent research and expert commentary at the most relevant time — not just in a yearly, themed issue. Please submit your best research, and feel free to contact us to discuss ideas for essays, viewpoints, editorials and commentaries.

And don't forget the Dr Ross Ingram Memorial Competition, open to any Aboriginal or Torres Strait Islander person who is working, researching or studying in a health-related field.

MJA

The Medical Journal of Australia

<http://www.mja.com.au/public/information/RossIngramPrize.html>

Contact: Dr Ruth Armstrong • mja@mja.com.au

