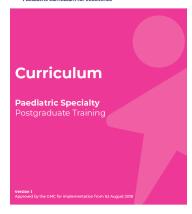


Rapid-fire critical appraisal for START assessments

Dr Ian Morris
Consultant
Neonatologist
University Hospital
of Wales

# Curriculum





The Royal College of Paediatrics and Child Health is a registered of



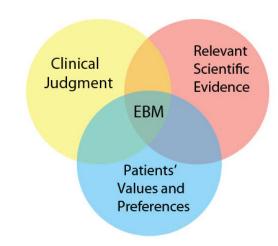
ı	Curriculum Domain	Level 1	Level 2	Level 3
1	1. Research	Adopts an evidence-	Implements an	Demonstrates
	GPC 9	based approach to paediatric health practice and the critical appraisal of existing published research.	evidence-based approach to practice to inform decision- making and enhance patient care and patient outcomes.	independent development and revision of guidelines and procedures to improve service delivery, centred around current clinical research and evidence-based
				healthcare.

# Learning objectives

- Understand how appraisal is relevant to practice
- Understand how and why it is included in START
- Develop a framework for rapid appraisal
- Understand basics of study design, bias and key terms

# Relevance to practice

- Making decisions
- Justifying decisions
  - Parents
  - Colleagues
  - Nursing staff
  - Yourself!
- Delivering high quality, safe medicine





### **START**

- ST7
- 12 x 8 minute stations (4 mins prep)
  - CBD type format
- 44 mins prep for prescribing and appraisal
- Standard: newly appointed consultant
- Form basis for PDP for ST8 year

- Dress smart
- Be confident and decisive

# **Key points**

#### Think about:

- Leadership
- Prioritisation
- Delegation
- Team working
- communication
- Learning opportunities for trainees
- Safeguarding
- Quality and safety of service



#### INFORMATION FOR TRAINEE

This is a 8 minute scenario. Check the 'Task' below so you know how this scenario will be run. You will have up to 4 minutes before the start of the assessment to read this information and prepare yourself. You may make notes on the paper provided.

#### You should

- Enter the assessment room when the bell sounds
- Take this instruction sheet with you and any notes you have prepared
- Note that there will be a warning when you have approximately 2-minutes left
- Return the instruction sheet outside the room when you leave

Context	You are a newly appointed cosultant		
Task Read the scenario and prepare an appraisal of the provided paper (Archives of			
	Disease in Childhood 2006; (5):414.		
Information for	A 3 year old with a history of a febrile convulsion at the age of 16 months arrives in the		
trainee	children's emergency department. He is bright and alert but has an obvious upper respiratory tract infection. His temperature is 38.7°C. His parents and the nursing staff are keen to give paracetamol and ibuprofen together to reduce the temperature as quickly possible. Is this known to be helpful? Review the paper and decide to what extent this supports or doesn't support the concurrent use of these medications.		

#### SHORT REPORT

### Randomised controlled trial of combined paracetamol and ibuprofen for fever

M D S Erlewyn-Lajeunesse, K Coppens, L P Hunt, P J Chinnick, P Davies, I M Higginson, J R Benger



Arch Dis Child 2006;91:414-416. doi: 10.1136/adc.2005.087874

A randomised open label study of the combined use of paracetamal and ibupraten to rapidly reduce fever is reported. The advantage of using both medications is less than half a degree centigrade in the first hour, and insufficient to warrant routine use.

very day parents and healthcare professionals treat febrile children using buprofen and paracetamol. The practice of giving both medicines simultaneously is widespread, but unsupported by evidence. Their use has been driven in part by a perceived need to prevent febrile convulsion, although evidence that antipyretics prevent such convulsions is also lacking. We assessed the short term effectiveness of a combined dose of paracetamol and ibuprofen in reducing childhood fever.

#### METHODS

We conducted an open label, three arm randomised trial in our inner city Children's Emergency Department. Participants received suspensions of paracetamol 15 mg/kg, ibuprofen 5 mg/kg, or both.

Consecutive children between 6 months and 10 years old attending with a fever of 38.0°C or more were included. Children were excluded if they had received either drug in the last six hours, were shocked, immunosuppressed, or had other known contraindications to either medicine (see box 1). Carers gave written informed consent.

The primary outcome measure was the child's temperature at one hour. Secondary outcomes included temperature at two hours and the time spent in the department. Too few children had data at two hours to allow meaningful comparison, as they had already been discharged home. Secondary outcome analysis of the time spent on the unit did not add to our findings and is not reported. Temperatures were measured using a tympanometric thermometer (Thermoscan, Braun Ltd, UK) at the time of admission, the time medication was given (TO), one hour later (T1), and two hours later (T2) if the child had not been discharged. Temperatures were measured in the presenting ear by a single reading according to the manufacturer's instructions. Patiful ears were avoided and normally a single observer would measure each child.

The sample size was calculated from pilot data collected using study methods, operators, and equipment. We judged a temperature difference of 1.0°C at one hour to be of clinical significance. To have an 80% chance of detecting this difference, at the two sided 5% level and including a 15% do not trate before one hour. 40 children per group were required.

The allocation sequence was block randomised and generated independently of the research team. Allocations were placed in sequentially numbered sealed opaque

envelopes. Staff working in the department used the next envelope in the sequence to allocate participants.

Mean temperatures were compared using a one way ANOVA with and without covariate adjustment for baseline temperature. Multiple comparisons were performed using Scheffe tests.

The Gloucestershire Research Ethics Committee approved the study.

#### RESULTS

A total of 123 children were randomised between October 2004 and January 2005 (fig 1). Baseline data were similar in all three groups, except that more children were admitted to hospital in the combined group (13/36) compared to the ibuprofen (3/35) and paracetamol groups (5/35). A notes review of admissions showed that one child from the combined group looked non-specifically unwell following the administration of study medication, during a rapid temperature drop from 39.5°C to 37.7°C in one hour. She was admitted for observation, recovered spontaneously, and was discharged three hours later. All other admissions were not related to new clinical events following the administration of study medication. One child in the paracetamol group received a dose of 27.8 mg/kg in error. The child did not suffer any adverse consequences from this overdose. There were no other adverse events

All children with data at TI (n=108) were included in the primary analysis on an intention to treat basis. There was a significant difference between the three groups overall at TI (p=0.023; table 1), which was unchanged by adjustment for

#### Box 1: Exclusion criteria

- Paracetamol or ibuprofen given in the previous six hours
- Severe or life threatening infection
- Suspected chicken pox
- Cellulitis or other spreading skin infection
- Known to be immunosuppressed
- Allergy to either ibuprofen or paracetamol
- Medicated with warfarin, heparin, or antihypertensives
- Symptoms of active gastrointestinal bleeding
- Known coagulopathy
- Acute joundice
- Likely dehydration, defined as more than four episodes of diarrhoea or vomiting in the previous 24 hours
- Asthma, defined as a need for regular "preventer" medication
- Chronic renal, liver, or cardiac failure

524 Parents did not consent Antipyretics in the last 6 hours 326 123 Ibuprofer 40 41 42 Missing data Missing data Missing data Refused Refused 37 35 36

baseline temperature. Pairwise comparisons showed a significant difference between the combined group and paracetamol alone (mean baseline adjusted difference at Ti 0.35°C; 95% CI 0.010 to 0.60; p = 0.028), but not between the combined group and  $\theta$ uprofen (0.25°C, 95% CI =0.01 to 0.50; p = 0.166). The difference between the ibuprofen and paracetamol groups was not statistically significant (0.10°C, 95% CI =0.01 to 0.36; p = 0.735).

#### DISCUSSION

Combined paracetamol and Buprofen were better at reducing fever after one hour than paracetamol alone. However, the effect is less than half a degree centigrade, and we do not believe that this is a clinically important difference over this time period.

We did not show a difference between combined therapy and ibuprofen. Our study was powered to detect a difference between combined therapy and paracetamol alone, and by doing so we have demonstrated proof of principle. The greater efficacy of ibuprofen' would require a larger sample size to show a significant difference.

This study was carried out in a paediatric emergency department, and therefore only examined the short term control of pyrexia. A longer measurement period might produce different results, as the maximum decrease in temperature for both medicines is around three hours

post-dose.\(^1\) In the emergency department we were unable to keep children back for study after discharge and in consequence, only one third (39/123) of children had their temperatures recorded two hours post-dose.

Figure 1 Flow thart of subject enrolment to one hour (T1) post-dose.

An alternative dosing schedule may also produce different findings. The synergism that we have demonstrated between paracetamol and ibuprofen suggests that alternating anti-pyretics may achieve better fever control over the course of an illness, but further research is required.

These medicines are frequently prescribed without incident, but occasionally have significant side effects including renal failure' and hypothermia. Thus the administration of both paracetamol and buprofen together should be used with caution.

In summary, although there is benefit from combined antipyretics, this effect is not large enough to warrant routine use for rapid fever reduction. In the context of paediatric emergency medicine, where rapid control of pyrexia is often desirable when making decision about disposition, we believe our findings will be useful when dealing with an acutely febrile child.

#### **ACKNOWLEDGEMENTS**

The authors would like to thank Samantha-Jayne Tanner and Jessica Blissett for their administrative help with this study. We would also like to thank the staff of the CED and the parents and children who participated.

	Paraceta mol	Ibuprofen	Both
Randomised	n=41	n=42	n=40
Age (years)	1.5 (0.6-9.5)	1.5 (0.5-9.6)	2.4 (0.6-8.2)
Weight (kg)	11.4 (7.0-47.0)	12.0 (7.5-33.0)	12.6 (7.9-25.0)
Dose of paracetamol (mg/kg)	15.3* (SD 2.0)	_	14.9 (SD 0.8)
Dose of ibuprofen (mg/kg)	-	5.0 (SD 0.2)	4.9 (SD 0.2)
[] data available (n)	n=37	n =35	n=36
Baseline (TO)	38.93 (SD 0.68)	38.73 (SD 0.63)	38.81 (SD 0.79)
One hour (TI)	37.98 (SD 0.47)	37.81 (SD 0.69)	37.59 (SD 0.61)
Mean fall from TO to T1 (95% CI)	0.95 (0.72-1.17)	0.92 (0.70-1.14)	1.22 (0.95-1.50)

#### Authors' affiliations

England, Bristol, UK

M D S Erlewyn-Lajeunesse, L P Hunt, Clinical Sciences at South Bristol, Faculty of Medicine, University of Bristol, UK

K Coppens, P J Chinnick, P Davies, I M Higginson, Children's Emergency Department, Bristal Royal Hospital for Children, Bristal, UK J R Benger, Faculty of Health and Social Care, University of the West of

Funding: this study was funded by the Anthony Hopkins Memorial Prize, awarded by the Faculty of Accident and Emergency Medicine as an unrestricted award to the Emergency Department

Competing interests: none dedared

Carrespondence to: Dr M Erlewyn-Lajeunesse, Institute of Child Life & Health, UBHT Education Centre, Upper Maudlin Street, Bristol BS2 8AE, UK: Mich. Lajeunesse @bristol.ac.uk

Accepted 1 February 2006 Published Online First 7 February 2006

#### REFERENCES

- Mayord CE, Marino RV, Rosenfeld W, et al. Alternating antipyretics: is this an alternative? Padiatrics 2000;105:1009-12.
- Uhari M, Rantala H, Vainionpaa L, et al. Effect of acetaminophen and of low intermittent doses of diazepam on prevention of recurrences of febrile seizures. J Pail pt 1995;126:991-5.
- 3 Perrott DA, Piira T, Goodenough B, et al. Efficacy and safety of acetaminophen vs ibuprofen for treating children's pain or fever: a metaanalysis. Arch Pediatr Adolesc Med 2004;158:521-6.
- 4 Moghal NE, Hegde S, Eastham KM. Ibuprofen and acute renal failure in a toddler. Arch Dis Child 2004;89:276-7.
- 5 Richardson J, Sills J. Hypothermia following fever. Arch Dis Child 2004;89:1177.

# Approach to appraisal

- Many different ways.....
- Many different checklists....
- 3 key areas:
  - Relevance
  - Internal validity
  - External validity

#### Relevance:

- Is this article important to me, my patients and my practice
- Title and abstract

#### Internal validity

- Soundness of the study
- Methods

#### **External validity**

- Applicability to my practice / generalisability
- Methods, results

BASIS FOR COMPARISON	INTERNAL VALIDITY	EXTERNAL VALIDITY
Meaning	Internal validity is the extent to which the experiment is free from errors and any difference in measurement is due to independent variable and nothing else.	External validity is the extent to which the research results can be inferred to world at large.
Concerned with	Control	Naturalness
What is it?	It is a measure of accuracy of the experiment.	It checks whether the casual relationship discovered in the experiment can be generalized or not.
Identifies	How strong the research methods are?	Can the outcome of the research be applied to the real world?
Describes	Degree to which the conclusion is warranted.	Degree to which the study is warranted to generalize the result to other context.
Used to	Address or eliminate alternative explanation for the result.	Generalize the outcome.

BASIS FOR COMPARISON	INTERNAL VALIDITY	EXTERNAL VALIDITY
Meaning	Internal validity is the extent to which the experiment is free from errors and any difference in measurement is due to independent variable and nothing else.	External validity is the extent to which the research results can be inferred to world at large.
Concerned with	Control	Naturalness
What is it?	It is a measure of accuracy of the experiment.	It checks whether the casual relationship discovered in the experiment can be generalized or not.
Identifies	How strong the research methods are?	Can the outcome of the research be applied to the real world?
Describes	Degree to which the conclusion is warranted.	Degree to which the study is warranted to generalize the result to other context.
Used to	Address or eliminate alternative explanation for the result.	Generalize the outcome.

# Paper analysis

#### 1) Describe the study

- What type of study (method)
- What type of question (diagnostic, therapeutic, economic)
- Where was it done (including single/multi-centre)
- Key feature (e.g. first of its kind)

Example: "This was a multicentre, randomised, controlled trial of therapy, and the latest of a number of studies assessing the use of probiotics to prevent necrotising enterocolitis in very preterm infants."

### 2) Describe the research question

- Population (who)
- Intervention (treatment, test)
- Comparison (control)
- Outcome

#### Example:

- P in infants born between 23 and 30 weeks' gestation
- I do daily probiotics (Bifodbacterium *breve* BBG-001)
- C versus placebo
- O reduce necrotising enterocolitis, late-onset sepsis, and death before discharge

- 3) State the relevance of this questionThis is usually described in the introduction
  - Correlate it to your clinical practice / clinical scenario
- 4) Describe the methods
  - Use PICO format but expand

#### Example:

- P-24 hospitals, SE England, 23+0-30+6
- excluded potentially lethal or GI malformations identified within 48 hours
  - excluded those no realistic chance of survival
    - IC Randomised within 48 hours of birth (web-based, NPEU)
  - Masking of parents, clinicians, outcome assessors
- OD B breve BBG-001 or placebo (identical packaging and preparation procedures)
  - started asap after randomisation, irrespective of enteral feeding
  - continued up to 36 weeks PMA / discharge
  - Stool cultures + 16s rRNA for B Breve at 2 weeks and 36 wk PMA
  - clinical data recorded

- O any episode NEC bell stage 2 or 3
- sepsis : any LOS (<46 PMA) with positive blood culture that is not skin commensal
  - death before discharge from hospital

#### secondary outcomes

number of cultures, cultures per infant, positive culture of B Breve from sterile site, antibiotic use, antifungal use, time to full enteral feeds, LOS, BPD, ROP......

#### Estimated 1300 infants needed

- 5% sig level, 90% power to detect 40% risk reduction in primary outcomes (15% to 9.1%)
  - ITT analysis

#### 5) Comment on internal validity

- Various checklists e.g. CONSORT, STROBE, PRISMA
- Is there a risk of bias?
- Were experiment and control groups similar at baseline?
   (randomisation, concealment, ITT)
- Did groups remain similar after study started? (blinding, follow-up)
- Attempt to minimise type 1 and type 2 errors?

## Bias

- A bias is a systematic error, or deviation from the truth, in results or inferences
- Multiple replications of the same study would reach the wrong answer on average

cf

- Random error or imprecision
- Quality of study

# Sources of Bias - 1

- Selection bias (randomisation)
  - systematic differences between baseline characteristics of the groups that are compared
  - sequence generation; allocation sequence concealment
- Performance bias (blinding of operators)
  - systematic differences between groups in the care that is provided, or in exposure to factors other than the interventions of interest
- Detection bias (blinding of assessors)
  - systematic differences between groups in how outcomes are determined

# Sources of Bias - 2

- Attrition bias (incomplete outcome data)
  - systematic differences between groups in withdrawals from a study (cf exclusions)
  - "intention to treat" analysis
- Reporting bias (selective reporting)
  - systematic differences between reported and unreported findings (analyses with statistically significant differences between intervention groups are more likely to be reported than nonsignificant differences – "within-study publication bias")

### Chance

Null hypothesis: Prediction that there will be NO significant difference in an outcome

Type I vs Type II error

	Null hypothesis $(H_0)$ is true	Null hypothesis $(H_0)$ is false
Reject null hypothesis	Type I error False positive	Correct outcome True positive
Fail to reject null hypothesis	Correct outcome True negative	Type II error False negative

In colloquial usage type I error can be thought of as "convicting an innocent person" and type II error "letting a guilty person go free".

- 6) Summarise the primary results
- 7) Summarise key secondary results
- 8) Describe the generalisability of these findings (external validity)
  - Context of your local population
  - Were inclusion / exclusion criteria reflective
  - Limitations that may impact
  - In conjunction with other studies

### 9) Conclude

Overall, does this paper help you answer your clinical question

# Worked example

#### Study

This was a single-centre, open label, randomised trial of therapy, assessing the use of antipyretics in reducing fever in children

#### **Research question**

#### Example:

P – in children between 6 months and 10 years with a fever

I – does combined ibuprofen and paracetamol

C – versus paracetamol or ibuprofen alone

O – reduce temperature at 1 hour

#### Relevance

Fever is a common presenting feature and cause for concern for parents and healthcare professionals alike. The use of multiple antipyretics is common, but unsupported by evidence. This paper thus seeks to answer an important and relevant question and is worth exploring.

#### Methodology

P – Children admitted to a UK Children's ED with a fever > 38°C were included. A number of exclusions are listed including those receiving antipyretics in the previous 6 hours and those with severe infection.

IC – Patients were randomised to receive either suspensions of paracetamol 15mg/kg, ibuprofen 5mg/kg or both in an open-label design.

O - The primary outcome was temperature at 1 hour as measured by a tympanometric thermometer.

#### **Internal validity**

A sample size calculation suggested 40 participants per group were needed for an 80% chance of detecting a 1°C difference at the 95% confidence level.

The groups were similar at baseline and were analysed on an intention to treat basis.

A high and equal number of children were analysed at 1 hour in each group.

Participants and assessors were not masked to intervention.

There is thus low risk of selection and attrition but high risk of performance and detection bias.

Overall, the validity of this study is reasonable.

#### **Key results**

In total, 123 children were randomised and 108 were included in primary analysis with an equal number across the 3 groups.

There was a statistically significant difference in temperature between the combined group and those receiving paracetamol alone. This was of 0.35°C mean reduction after adjusting for baseline variables.

No other differences in primary or secondary outcomes were seen.

#### Generalisability

Patients included in this study were similar to the population we would expect locally, with broad inclusion criteria. The observation time was in line with what would be relevant to our department. No comment was made on other clinical observations nor on important history points such as a history of febrile convulsions or source of fever.

#### **Conclusion**

This well designed, small study failed to show a significant clinical benefit of combined antipyretics for reducing fever at 1 hour in febrile children. It provides further evidence for a more conservative use of antipyretics, but I would consider results of similar studies before adjusting my practice.

# Concepts to understand

- Study design
  - RCT, Observational studies, systematic reviews
  - Randomisation, blinding, intention to treat
- Chance
  - Type 1 and type 2 errors
  - Sample size calculations
- Bias
- Statistical Vs Clinical significance
- Basic statistics
  - Key types of test
  - P value
  - Means, confidence intervals
  - RR, OR, NNT, NNH, PPV, NPV

# Sample size

• Importance: waste time, waste resources, ethics

• 4 components:

- Type I error ( $\alpha$ ) Usually 5%

Type II error (β)Usually 10 or 20%

- Variability ( $\sigma^2$ ) e.g. standard deviation

Effect size (d)
 Difference that would be clinically relevant

- Standard formulas, e.g N =  $16\sigma^2 / d^2$
- Sample size increases if variance increases or effect size decreases

# Common tests

#### Parametric:

assumes a form (usually normal) of distribution of data

#### Non-parametric:

no assumptions
looks at rank order of
values, ignores absolute
differences

Makes showing significance harder

lable	Some	commonly	useu	statistice	II LCS	LS
				_		_

Parametric test	Example of equivalent non-parametric test	Purpose of test	Example	
Two sample (unpaired) t test	samples drawn from the same		To compare girls' heights with boys' heights	
One sample (paired) ttest	Wilcoxon matched pairs test	Compares two sets of observations on a single sample	To compare weight of infants before and after a feed	
One way analysis of variance (Ftest) using total sum of squares	Kruskall-Wallis analysis of variance by ranks	Effectively, a generalisation of the paired t or Wilcoxon matched pairs test where three or more sets of observations are made on a single sample	To determine whether plasma glucose level is higher one hour, two hours, or three hours after a meal	
Two way analysis of variance	Two way analysis of variance by ranks	As above, but tests the influence (and interaction) of two different covariates	In the above example, to determine if the results differ in male and female subjects	
$\chi^2$ test	Fisher's exact test	Tests the null hypothesis that the distribution of a discontinuous variable is the same in two (or more) independent samples	To assess whether acceptance into medical school is more likely if the applicant was born in Britain	
Product moment correlation coefficient (Pearson's <i>t</i> )	Spearman's rank correlation coefficient (r183 $\sigma$ )	Assesses the strength of the straight line association between two continuous variables.	To assess whether and to what extent plasma HbA1 concentration is related to plasma triglyceride concentration in diabetic patients	
Regression by least squares method	regression quantitative va		To see how peak expiratory flow rate varies with height	
Multiple regression by least squares method	Non-parametric regression (various tests)	Describes the numerical relation between a dependent variable and several predictor variables (covariates)	To determine whether and to what extent a person's age, body fat, and sodium intake determine their blood pressure	

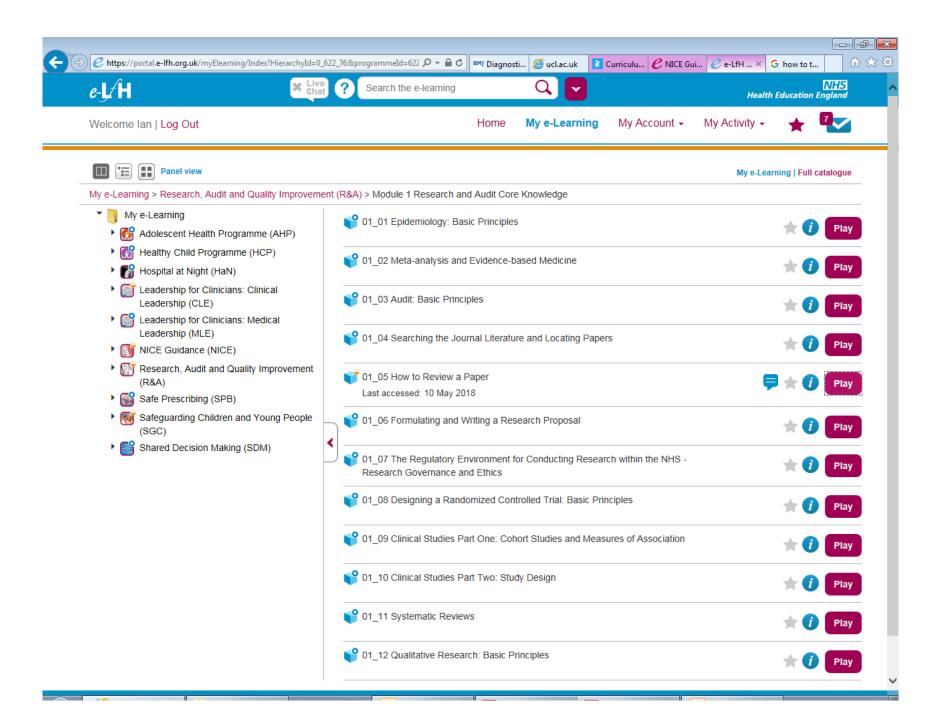
# Resources

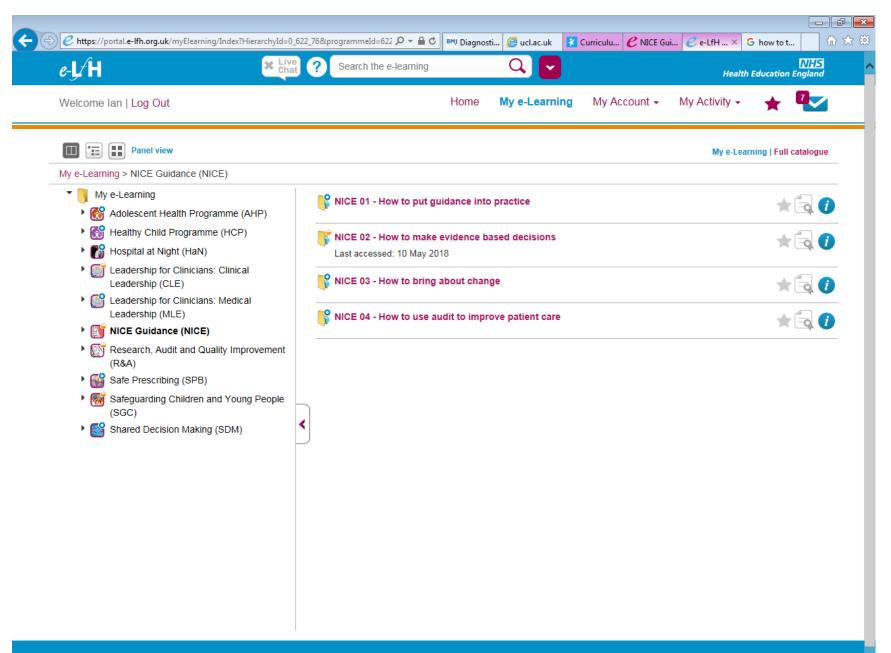
 How to read a paper: Statistics for the nonstatistician. I: Different types of data need different statistical tests

How to read a paper: Statistics for the nonstatistician. II: "Significant" relations and their pitfalls BMJ 1997; 315 Greenhalgh

- https://www.bmj.com/about-bmj/resourcesreaders/publications/statistics-square- one
- https://lifeinthefastlane.com
- http://training.cochrane.org
- E-learning for health, <a href="https://portal.e-lfh.org.uk/">https://portal.e-lfh.org.uk/</a>







 Ability to appraise and apply EBM crucial

# Summary

- Develop a simple, logical approach
  - Relevance
  - Internal validity
  - External validity
- Don't get bogged down with statistics

 Build skills through free courses and work place practice

# Questions?

Thank you

lan.morris3@wales.nhs.uk