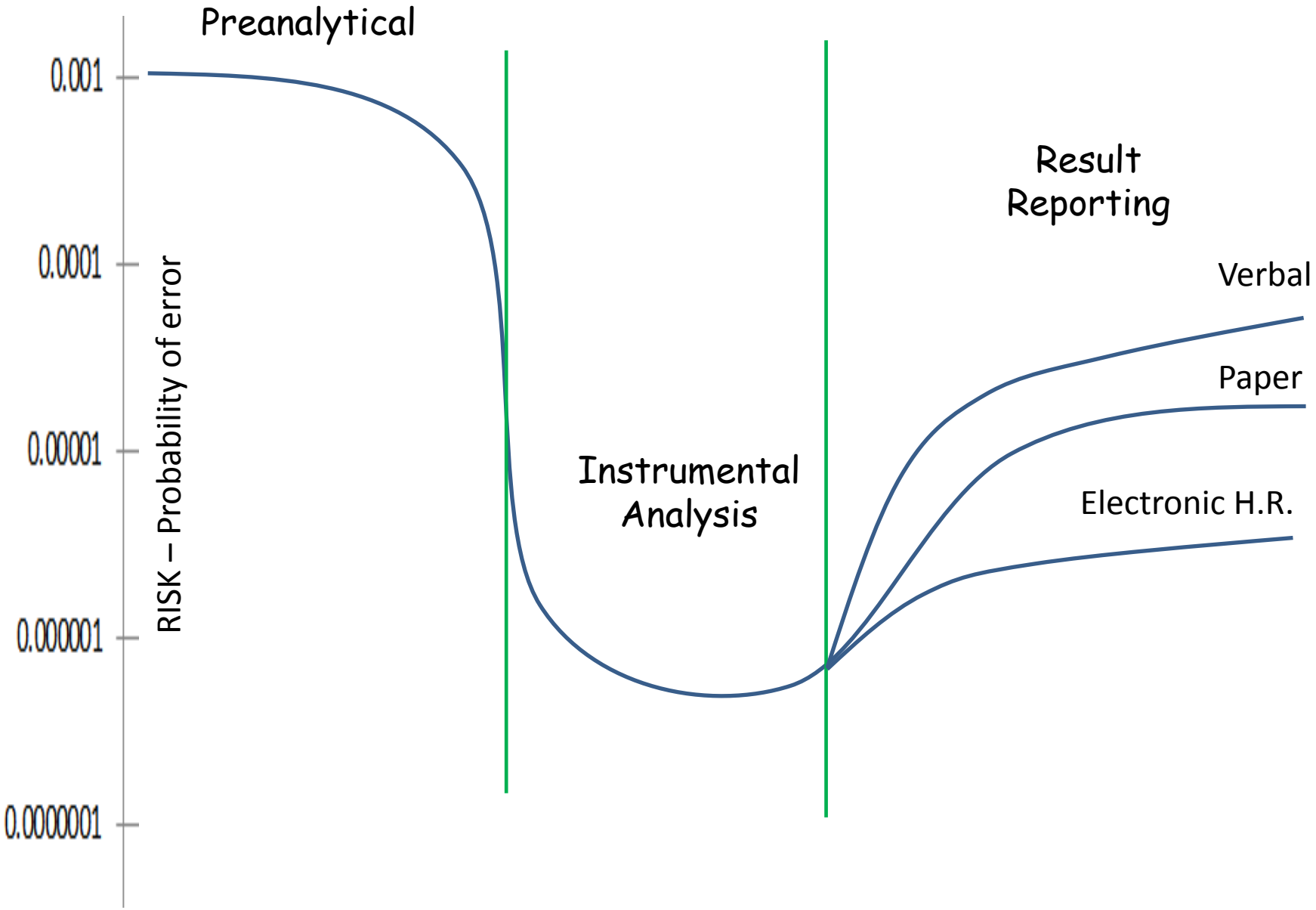
GETTING "REAL DATA" TO DRIVE QUALITY IMPROVEMENT AND REDUCE "REAL RISK" IN HEALTHCARE

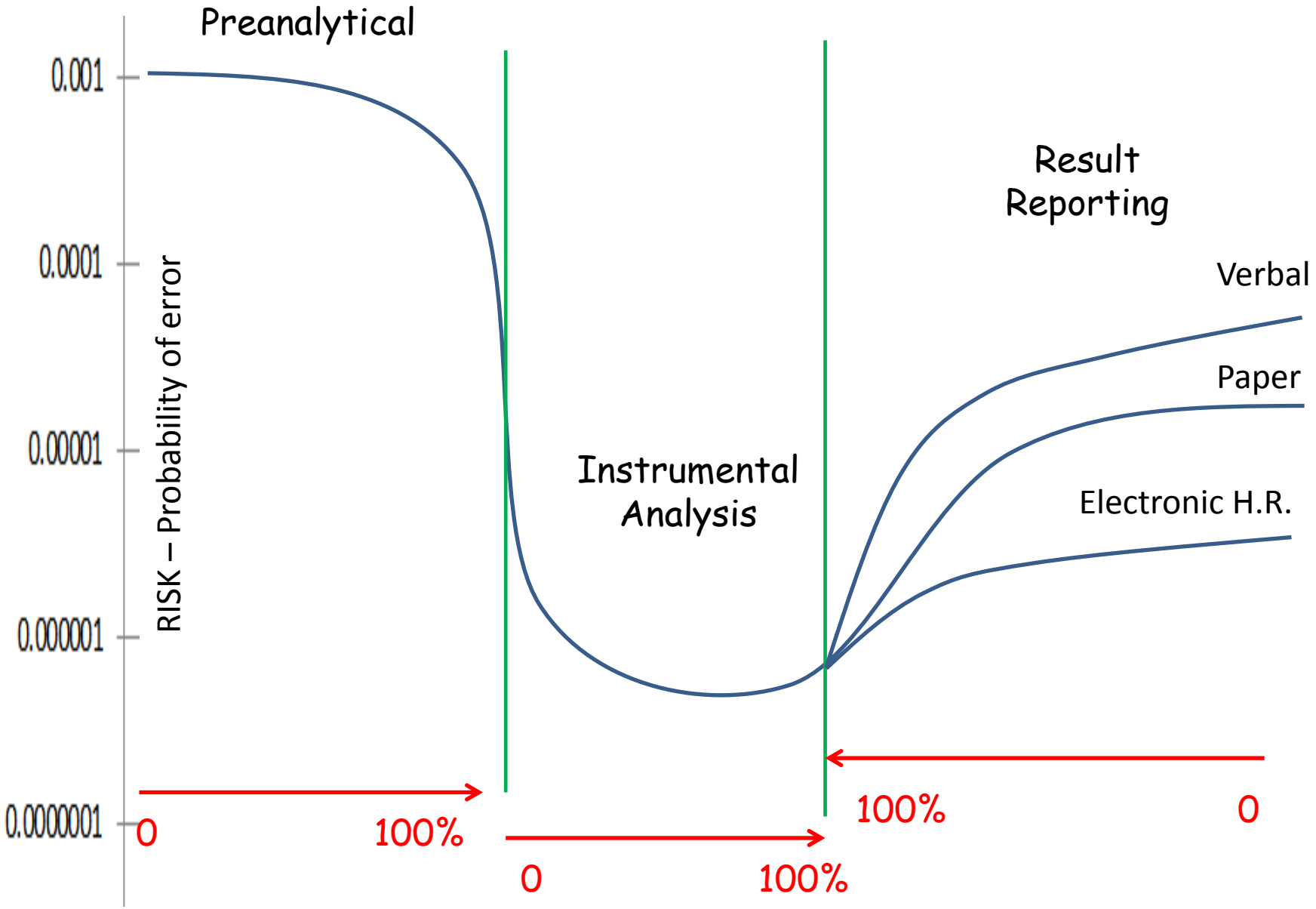
Dr Tom Hartley
Quality Manager
RHH Pathology Services
March 2013

MY THEME : WITHOUT DATA YOU ARE JUST ANOTHER PERSON WITH AN OPINION

PROPOSED RISK PROFILE ACROSS THE THREE STAGES OF MEDICAL TESTING

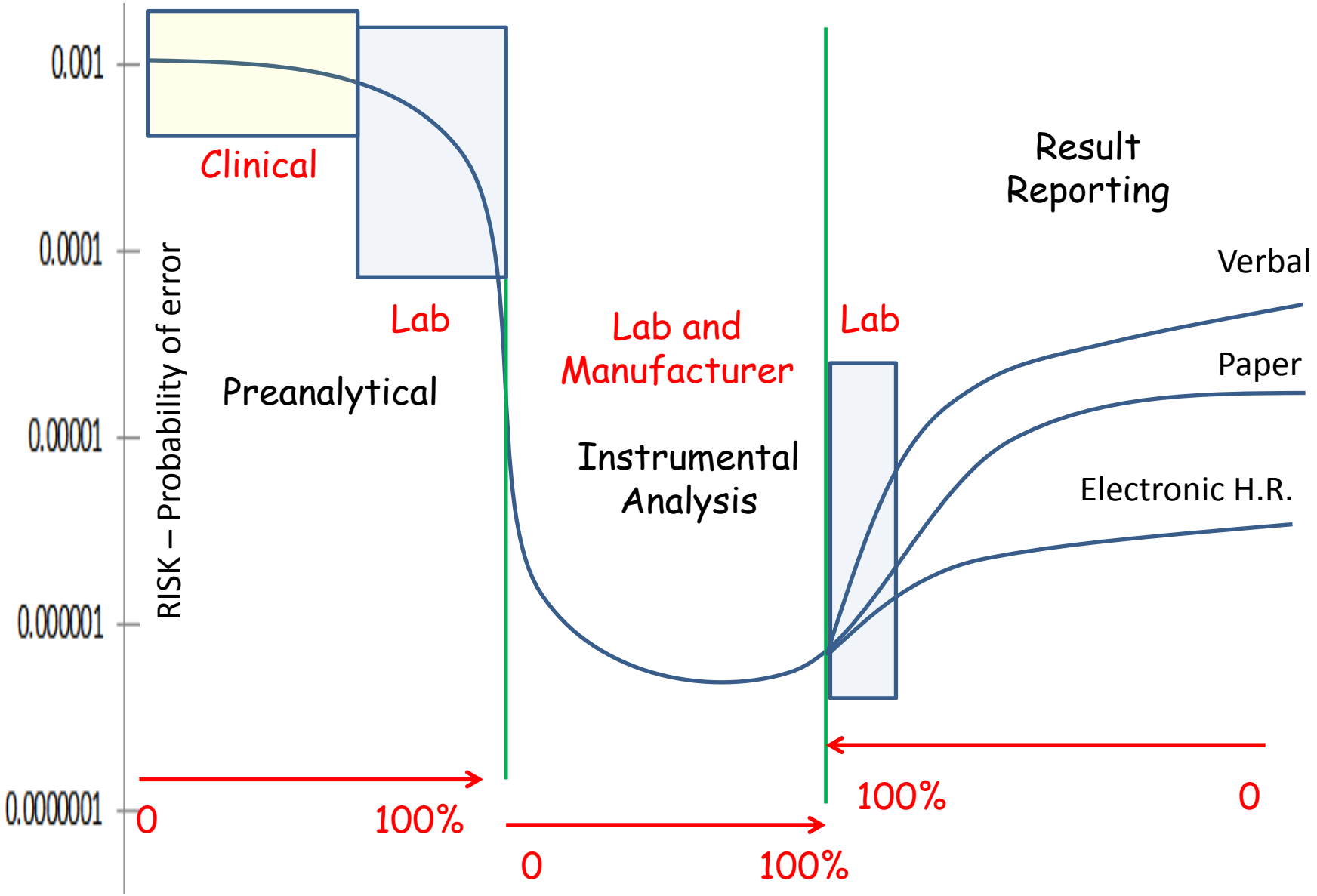


PROPOSED RISK PROFILE ACROSS THE THREE STAGES OF QUANTITATIVE MEDICAL TESTING



OPERATOR'S COMPETANCE AND VIGILANCE

RISK 'OWNERS' IN THE THREE STAGES OF QUANTITATIVE MEDICAL TESTING



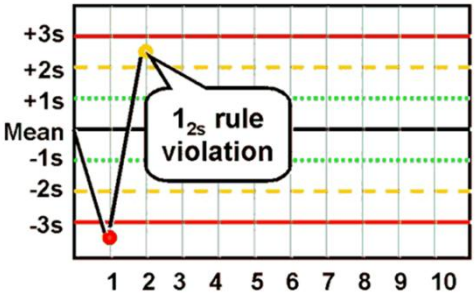
OPERATOR'S COMPETANCE AND VIGILANCE

SO WHERE IS THE DATA TO SUPPORT THIS PROPOSED RISK PROFILE ACROSS THE THREE STAGES OF MEDICAL TESTING ?

RISK OF ANALYTICAL ERROR IN QUANTITATIVE TESTING

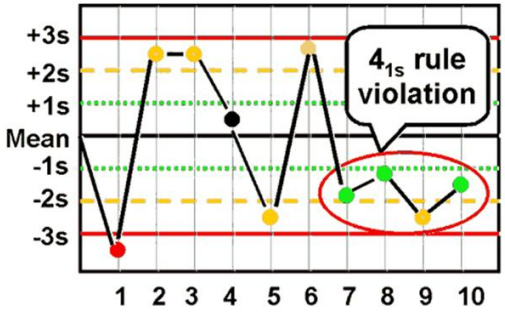
If you use a ‘kit’ from a major manufacturer and use the Westgard Rules then the risks of error are :

p = 0.0013



p = 0.0228

p = 0.0228^2 = 0.0005



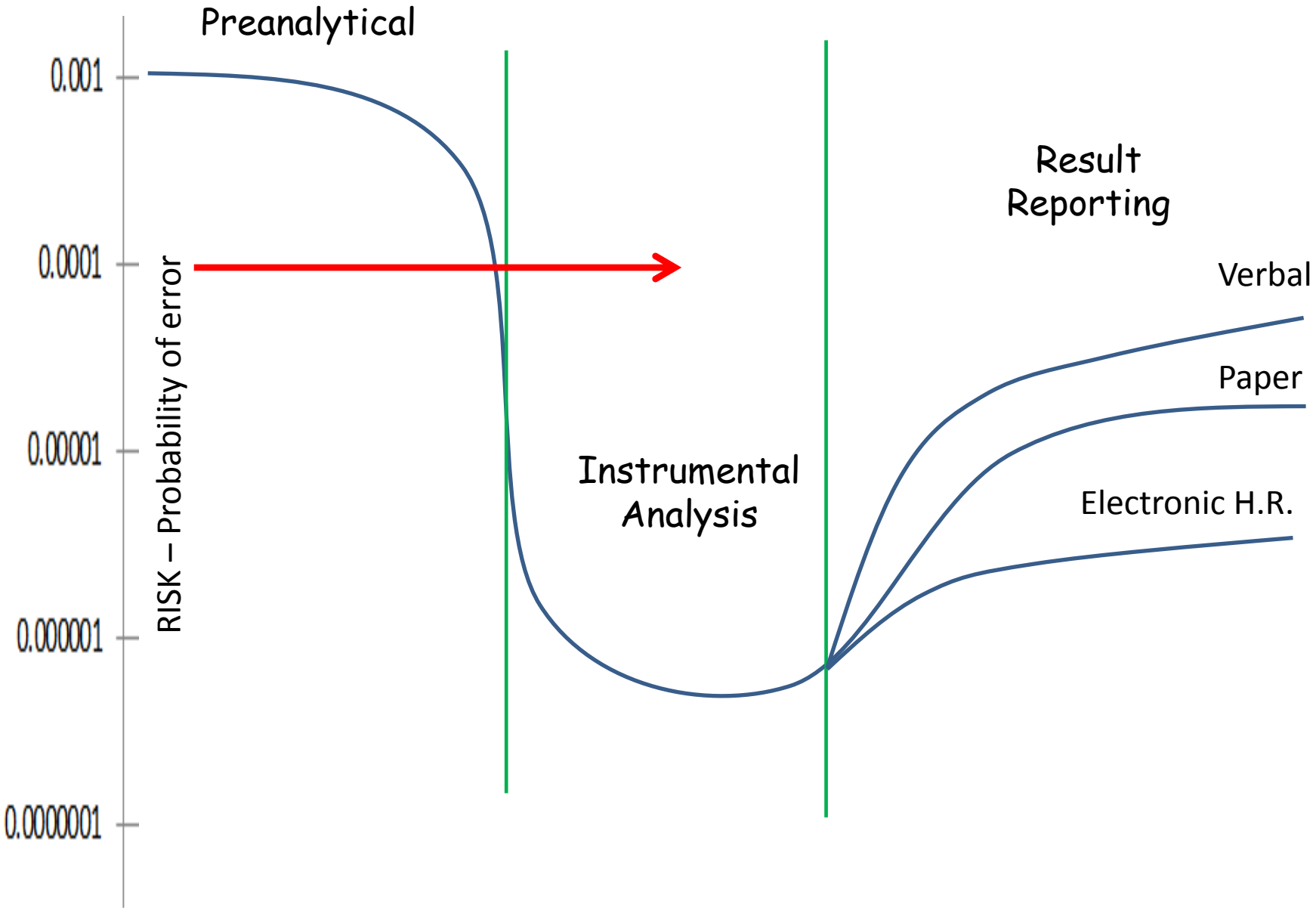
p = 0.0006

MEDIAN RISK OF ERROR IN QUANTITATIVE MEASUREMENT

p = 0.00095

≈ 0.0001

PROPOSED RISK PROFILE ACROSS THE THREE STAGES OF MEDICAL TESTING



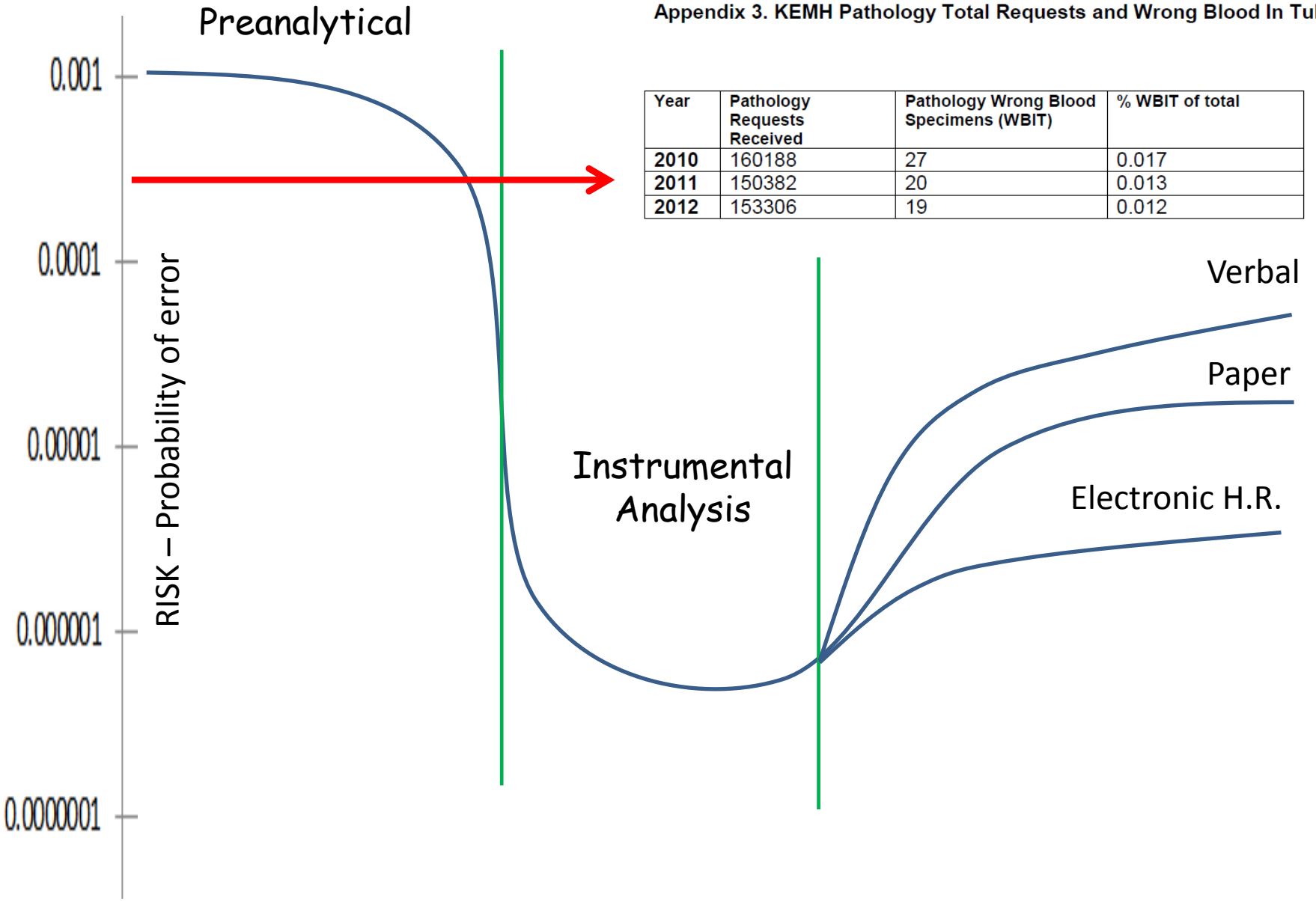
Risk of Preanalytical Errors Data from King Edward Memorial Hospital, Perth, WA

Appendix 3. KEMH Pathology Total Requests and Wrong Blood In Tube

Year	Pathology Requests Received	Pathology Wrong Blood Specimens (WBIT)	% WBIT of total
2010	160188	27	0.017
2011	150382	20	0.013
2012	153306	19	0.012

p = 0.00012 to 0.00017

PROPOSED RISK PROFILE ACROSS THE THREE STAGES OF MEDICAL TESTING



Appendix 3. KEMH Pathology Total Requests and Wrong Blood In Tube

Year	Pathology Requests Received	Pathology Wrong Blood Specimens (WBIT)	% WBIT of total
2010	160188	27	0.017
2011	150382	20	0.013
2012	153306	19	0.012

KIMMS 2012 SUMMARY
Preanalytical – part 1

Statistics Summary KIMMS 2012	Jan-Mar			Apr-Jun			Jul-Sept			Oct-Dec		
PRE-ANALYTICAL	All (68)			All (70)			All (67)			All (70)		
IDENTIFICATION PROBLEMS	Count	All %	% of Accessions	Count	All %	% of Accessions	Count	All %	% of Accessions	Count	All %	% of Accessions
Sample suspected to be from wrong patient (wrong patients blood in tube)	455	2.38	0.01	476	2.32	0.01	360	1.61	0.01	379	1.76	0
Unlabelled samples	5098	26.68	0.08	5427	26.43	0.07	6623	29.88	0.11	5819	27.07	0.07
Fewer than 2 identifiers initially supplied	2196	11.49	0.03	3210	15.63	0.04	2817	12.62	0.05	3128	14.55	0.04
Any mismatch or discrepancy of identifiers (major or minor)	4318	22.6	0.07	4670	22.74	0.06	5260	23.57	0.09	5425	25.23	0.06
Any within laboratory failure of ID	920	4.81	0.01	762	3.71	0.01	799	3.58	0.01	738	3.43	0.01
Transfusion issues-not covered in other categories	3634	19.02	0.06	3506	17.07	0.05	3856	17.28	0.06	3971	18.47	0.05
Sample misidentifications not classified above	1293	6.77	0.02	1482	7.22	0.02	1522	6.82	0.02	1179	5.48	0.01
ID Errors – e-requests	706	3.69	0.01	577	2.81	0.01	625	2.8	0.01	410	1.91	0
Precious sample ID errors	487	2.55	0.01	424	2.06	0.01	453	2.03	0.01	449	2.09	0.01
TOTAL ID errors	19107			20534			22315			21498		
IDENTIFICATION errors as % of ACCESSIONS			0.29%			0.27%			0.36%			0.25%

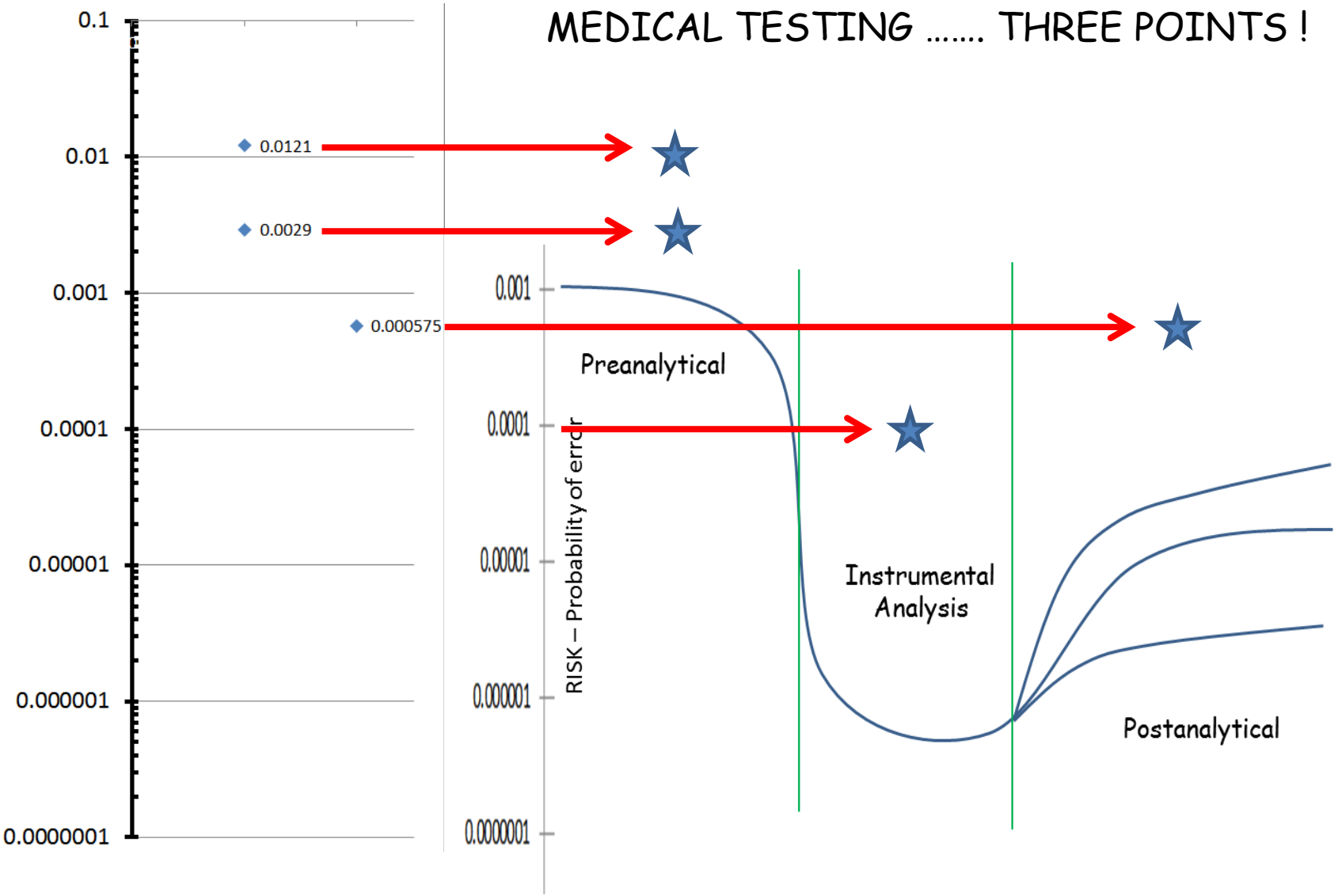
KIMMS 2012 SUMMARY
Preanalytical – part 2

Statistics Summary KIMMS 2012	Jan-Mar			Apr-Jun			Jul-Sept			Oct-Dec		
PRE-ANALYTICAL	All (68)			All (70)			All (67)			All (70)		
SAMPLES REJECTED	Count	All %	% of Accessions	Count	All %	% of Accessions	Count	All %	% of Accessions	Count	All %	% of Accessions
Samples rejected due to misidentification issues	7341	**8.47**	0.11	8150	**9.5**	0.11	7975	**9.65**	0.13	6980	**7.82**	0.08
Sample haemolysed	24432	**28.21**	0.38	22134	**25.79**	0.3	21812	**26.4**	0.36	22192	**24.86**	0.26
Sample clotted	7554	8.72	0.12	7770	9.05	0.1	7167	8.67	0.12	7746	8.68	0.09
Incorrect fill level of sample	3745	4.32	0.06	3861	4.5	0.05	5185	6.27	0.08	3340	3.74	0.04
Insufficient sample	7133	8.23	0.11	7319	8.53	0.1	5450	6.6	0.09	7789	8.73	0.09
Incorrect sample storage or transport	1843	2.13	0.03	1341	1.56	0.02	1485	1.8	0.02	2069	2.32	0.02
Specimen not collected	18860	**21.77**	0.29	18738	**21.84**	0.25	19232	**23.27**	0.31	20571	**23.04**	0.24
Incorrect specimen type	4202	4.85	0.06	4335	5.05	0.06	3813	4.61	0.06	4170	4.67	0.05
Registration of test error	3452	3.99	0.05	5244	6.11	0.07	3899	4.27	0.06	7335	8.22	0.08
Laboratory Accident / Error	3380	3.9	0.05	1581	1.84	0.02	1490	1.8	0.02	2360	2.64	0.03
Contaminated Sample	603	0.7	0.01	591	0.69	0.01	557	0.67	0.01	538	0.6	0.01
Precious samples rejected	110	0.13	0	291	0.34	0	424	0.51	0.01	347	0.39	0
Other (please specify)	3966	4.58	0.06	4458	5.2	0.06	4144	5.01	0.07	3831	4.29	0.04
TOTAL SAMPLE REJECTIONS	86621			85813			82633			89268		
REJECTIONS as % of ACCESSIONS			1.33%			1.15%			1.35%			1.03%

KIMMS 2012 SUMMARY
Postanalytical

POST ANALYTICAL 2012	Jan-Mar			Apr-Jun			Jul-Sept			Oct-Dec		
RESULTS CORRECTIED / REPORTS RETRACTED	All (68)			All (70)			All (67)			All (70)		
Report retracted because of an error after release in any form by the laboratory	3088	77.75	0.05	2721	70.42	0.04	3038	75.52	0.05	3071	71.64	0.04
Results released to wrong doctor	878	22.25	0.01	1143	28.58	0.02	985	24.48	0.02	1216	28.36	0.01
Root cause of post-analytical issues, if known	35			77						33		
TOTAL	3946			3864			4023			4287		
RESULT RETRACTION or CORRECTION as Percentage of ACCESSIONS			0.06%			0.05%			0.07%			0.05%

DATA TO SUPPORT THE PROPOSED RISK PROFILE ACROSS THE THREE STAGES OF MEDICAL TESTING ……. THREE POINTS !



SO IT LOOKS LIKE THESE EVENTS ARE NOT A RARE AS WE THOUGHT ? ……

Kaiser Health News

Is New US Patient Safety Effort Working?

Michael L. Millenson

Mar 01, 2012

For context, it helps to understand that the most widely quoted estimate of preventable patient harm — 44,000 to 98,000 deaths and one million injuries annually — was probably low. That estimate caused an uproar in a 1999 Institute of Medicine (IOM) report. Today, it seems conservative. The IOM total was based on studies conducted in hospitals in the mid-1980s. Recent research by the HHS Office of the Inspector General (OIG) and others has found a much higher rate of harm.

A Medicare patient today has a one-in-seven chance of suffering harm in the hospital, *(p = 0.14)*

a risk about four-to-seven times greater than in the IOM report. Moreover*,*

nearly 9 out of 10 incidents are never reported, the OIG concluded, even including incidents that led to patient deaths.

"If you measure all-cause harm, you find it in

about one-third of patients," (p=0.33)

says the University of Utah's Dr. David Classen, lead author of a 2011 study that appeared in Health Affairs.

USA population = 313,914,040

therefore 'Risk of Injury' in a USA hospital = 1 million/313 million (p= 0.003)

or 'Risk of harm leading to death' = 0.098 million/313 million (p = 0.0003)

WHAT IS LIMITING OUR ABILITY TO COLLECT DATA ?

1 : THE RELIANCE UPON PERSONAL COMPETANCE AND VIGILANCE.

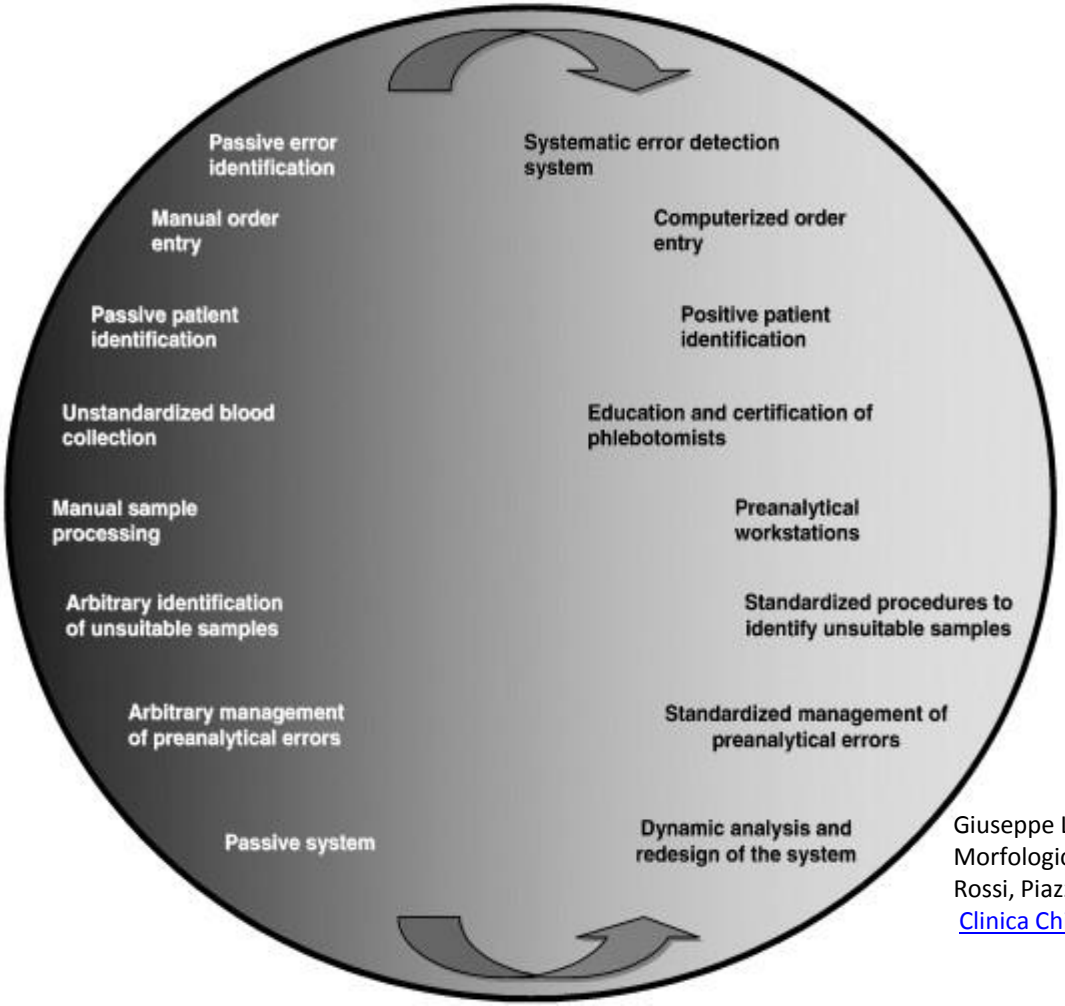
2 : THE STATISTICS OF RARE EVENTS IS THE INSURMOUNTABLE LIMIT.

WHAT IS LIMITING OUR ABILITY TO COLLECT DATA ?

THE RELIANCE UPON PERSONAL COMPETANCE AND VIGILANCE IS ONE MODIFIABLE LIMIT. THE MODIFICATIONS INCLUDE ADOPTION OF ‘FAIL SAFE’ PROCEDURES AND GREATER ‘IT’ MONITORING OF CRITICAL ACTIVITIES

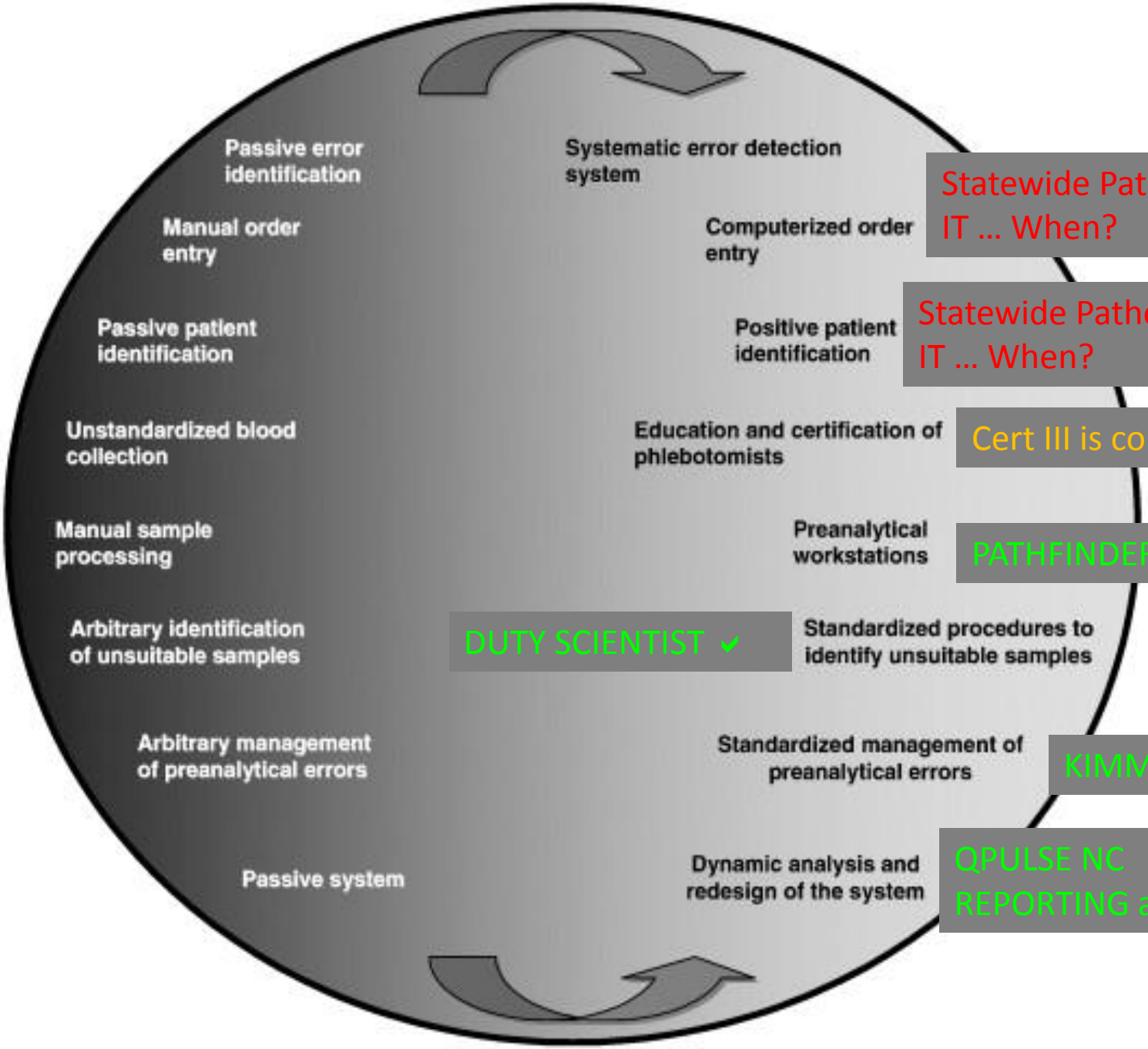
THE STATISTICS OF RARE EVENTS IS THE INSURMOUNTABLE LIMIT.

In our ‘backyard’ OLD WAYS must give way to NEW WAYS

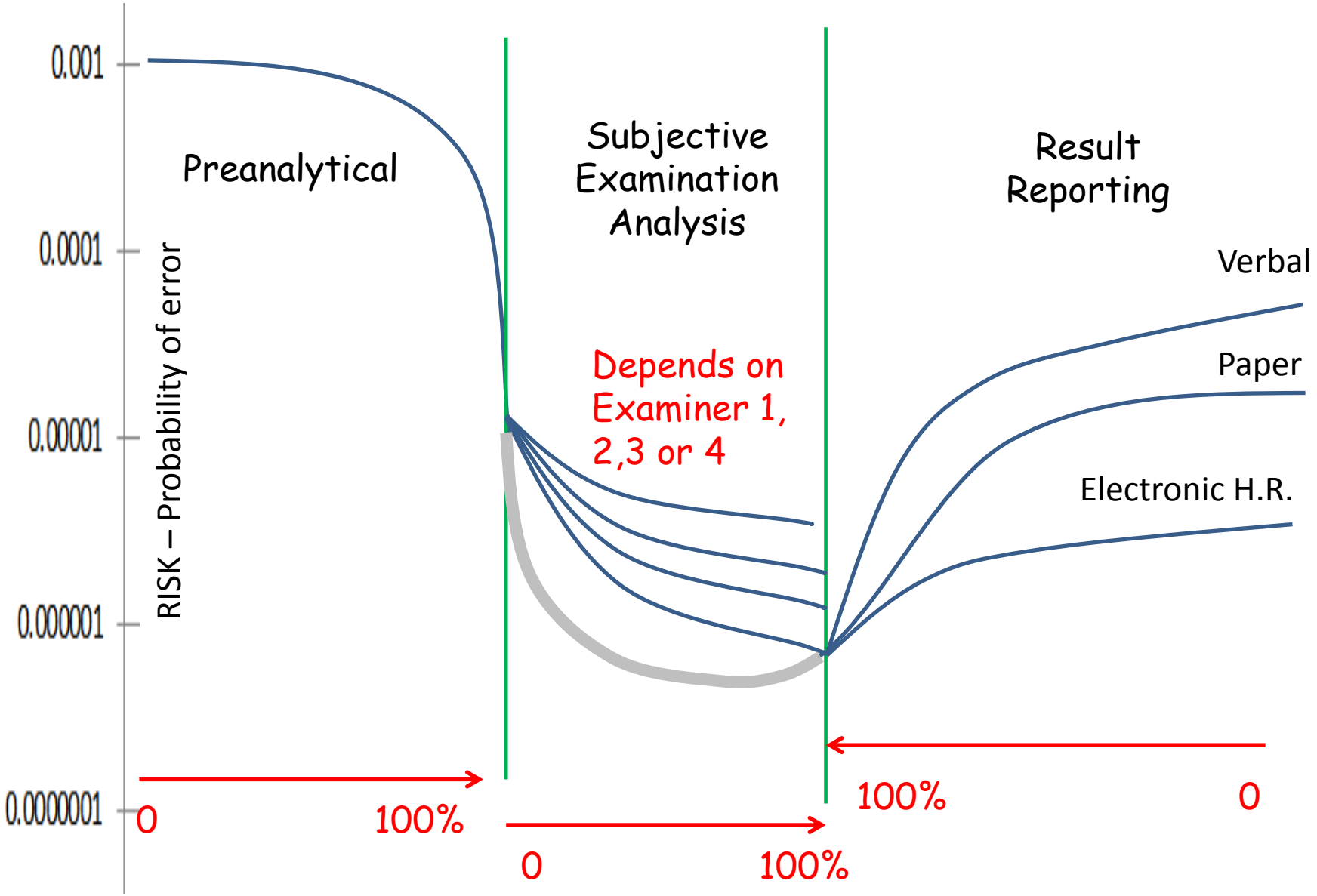


Governance of preanalytical variability: Travelling the right path to the bright side of the moon?

Giuseppe Lippi : Istituto di Chimica e Microscopia Clinica, Dipartimento di Scienze Morfologico-Biomediche, Università degli Studi di Verona, Ospedale Policlinico G.B. Rossi, Piazzale Scuro, 10, 37134-Verona, Italy
Clinica Chimica Acta Volume 404, Issue 1, 6 June 2009, Pages 32–36


Passive error identification
Systematic error detection system
Manual order entry
Computerized order entry
Passive patient identification
Positive patient identification
Unstandardized blood collection
Education and certification of phlebotomists
Manual sample processing
Preanalytical workstations
Arbitrary identification of unsuitable samples
Standardized procedures to identify unsuitable samples
Arbitrary management of preanalytical errors
Standardized management of preanalytical errors
Passive system
Dynamic analysis and redesign of the system
Statewide Pathology IT … When?
Statewide Pathology IT … When?
Cert III is coming
PATHFINDER ✔
DUTY SCIENTIST ✔
KIMMS ✔
QPULSE NC REPORTING and CA ✔

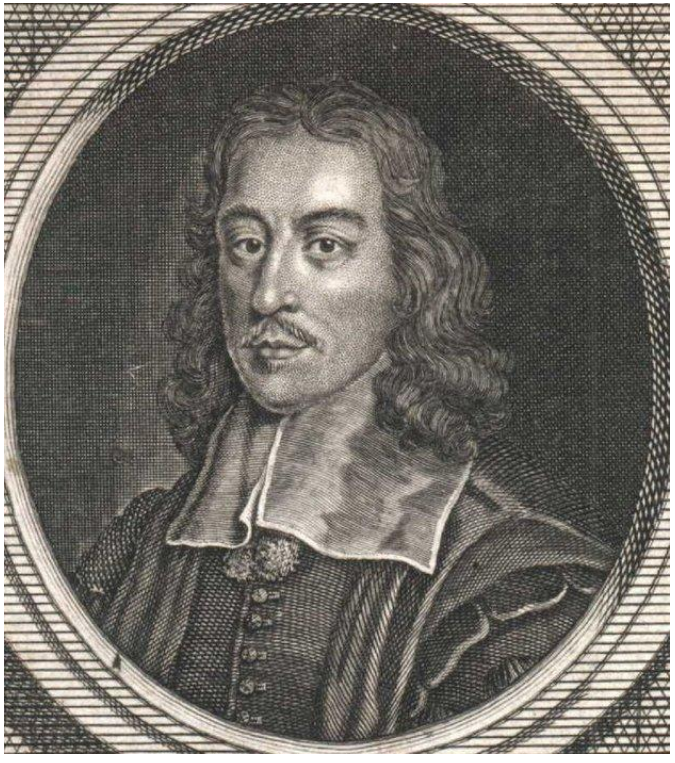
PROPOSED RISK PROFILE ACROSS THE THREE STAGES OF SUBJECTIVE MEDICAL TESTING



OPERATOR'S COMPETANCE AND VIGILANCE

SUBJECTIVE PATHOLOGY TESTING

1674 :English Physician, Thomas Willis, Coins the Term 'Diabetes Mellitus' due to the Sweetness of Diabetic Urine



2013

Introduction to NATA Tech Note #17 : Subjective Testing
What does a Qualitative Medical Test involve ?

- Sample preparation
- Placement in front of an expert
- Expert responds with an output opinion
- That opinion is compared with experience from similar specimens
- A probability is calculated
- That probability along with an interpretative guide is printed onto a report

How To Assess Microscopists Objectively ?

Suppose we want to assess a microscopist's performance we could do this by presenting them with 500 slides from a reference slide set for which there are 'expert consensus' classifications as shown in the Table. We could then look at how that microscopist performed versus the 'expert classifications'

GRADE	MICROSCOPIST 'A'	CONSENSUS NUMBER IN THIS GRADE
NORMAL	88	100
GRADE 1	122	100
GRADE 2	75	100
GRADE 3	110	100
GRADE 4	105	100
TOTALS	500	500

How To Assess Microscopists Objectively ?

Use the Chi Squares Test

GRADE	MICROSCOPIST 'A'	CONSENSUS NUMBER IN THIS GRADE	Use the CHI^2 test : Sum Of (Observed - Expected)^2 Divided by Expected
NORMAL	88	100	1.44
GRADE 1	122	100	4.84
GRADE 2	75	100	6.25
GRADE 3	110	100	1
GRADE 4	105	100	0.25
TOTALS	500	500	

TOTAL = Chi Squared	13.78
Degrees of Freedom	4
p =	0.008

Any value of Chi Squared greater than 9.49 would have been significant. In other words we can be 95% certain that the microscopist was deviating from the expert consensus

NATA Tech Note 17 : Subjective Testing

But the Chi Squared approach does not satisfy the requirements in : Section 5.2

- Probability of Detection
- Potential Error Rates – there are two
  the *False Negative Rate*
  and
  the *False Positive Rate*)

These can be derived from our microscopists study without any modification provided we look at the data more closely and use the formulae given on

page 12 of TN #17

Analyze the Microscopist's slide classifications ….

		CONSENSUS GRADING				
		NORMAL	GRADE 1	GRADE 2	GRADE 3	GRADE 4
MICROSCOPIST	NORMAL	80	5	3		
	GRADE 1	20	90	12		
	GRADE 2		5	70		
	GRADE 3			15	90	5
	GRADE 4				10	95
	TOTALS	100	100	100	100	100

		CONSENSUS GRADING				
		NORMAL	GRADE 1	GRADE 2	GRADE 3	GRADE 4
MICROSCOPIST	NORMAL	80	5	3		
	GRADE 1	20	90	12		
	GRADE 2		5	70		
	GRADE 3			15	90	5
	GRADE 4				10	95
	TOTALS	100	100	100	100	100

"DOWNGRADES"	FALSE NEGATIVES		5+3+12+5	**25**
"UPGRADES"	FALSE POSITIVES		20+5+15+10	**50**
	TRUE NEGATIVES		80	**80**
	TRUE POSITIVES		90+70+90+95	**345**

"DOWNGRADES"	FALSE NEGATIVES		5+3+12+5	25
"UPGRADES"	FALSE POSITIVES		20+5+15+10	50
	TRUE NEGATIVES		80	80
	TRUE POSITIVES		90+70+90+95	345
	SENSITIVITY	TP/(TP + FN)	0.93	
	SPECIFICITY	TN/(FP + TN)	0.62	
	POS PRED VALUE	TP/(TP + FP)	0.87	
	NEG PRED VALUE	TN/(FN + TN)	0.76	
	POD % = TP rate	100 * TP/Total #	69.0	
	PER = FP Rate	100 * FP / (TN + FP)	38.5	
	PER = FN Rate	100 * FN / (TP + FN)	6.8	

RISK OF ERROR IN QUALITATIVE TESTING

HERE ARE FIVE TISSUE SLIDES AND FIVE HAEMATOLOGY SLIDES – WHAT ARE THEY ?

THE POSSIBILITIES ARE

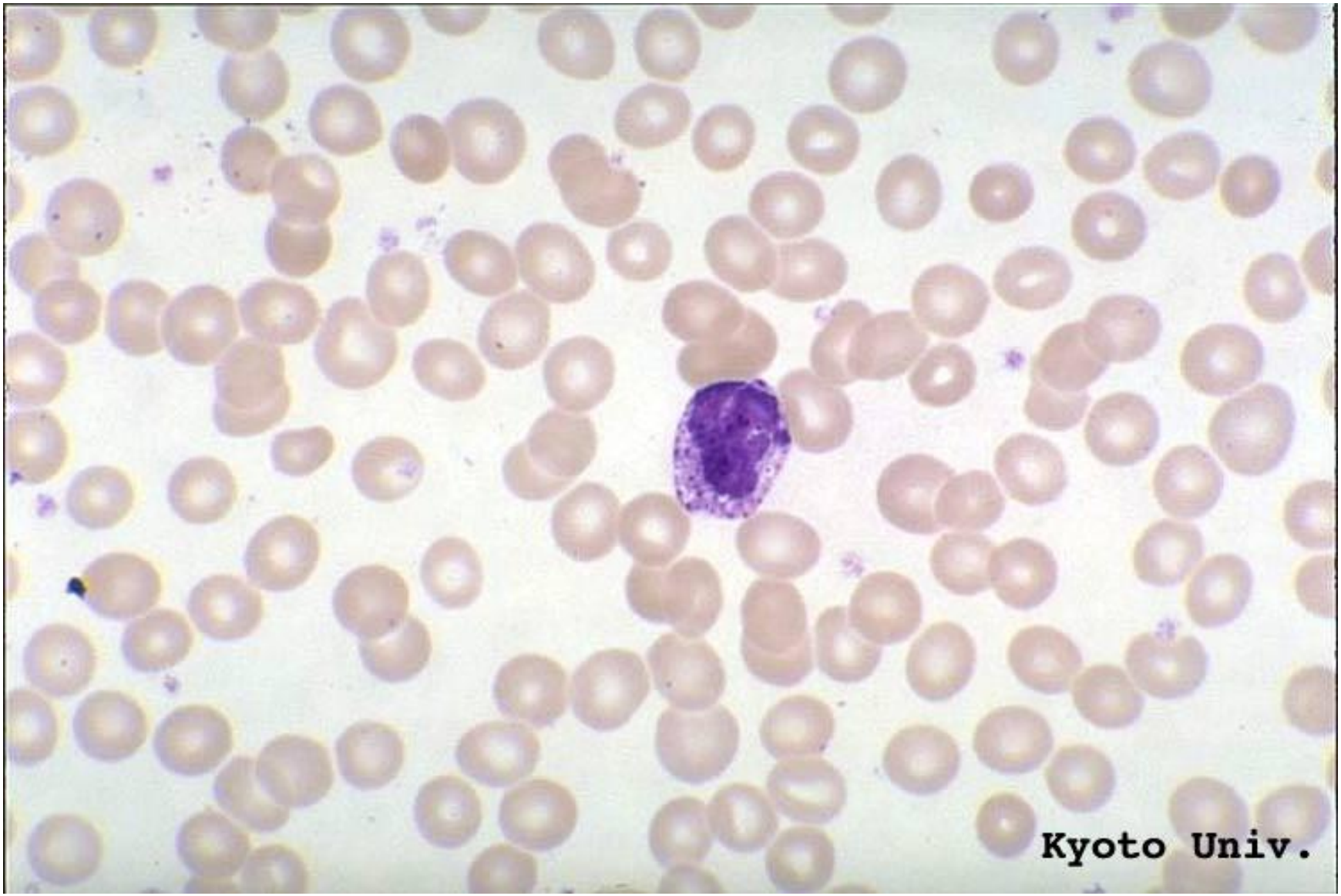
PANCREAS : LIVER : KIDNEY : LUNG : TESTES :
NEUTROPHIL : EOSINOPHIL : BASOPHIL : MONOCYTE

Record your answers on the card :

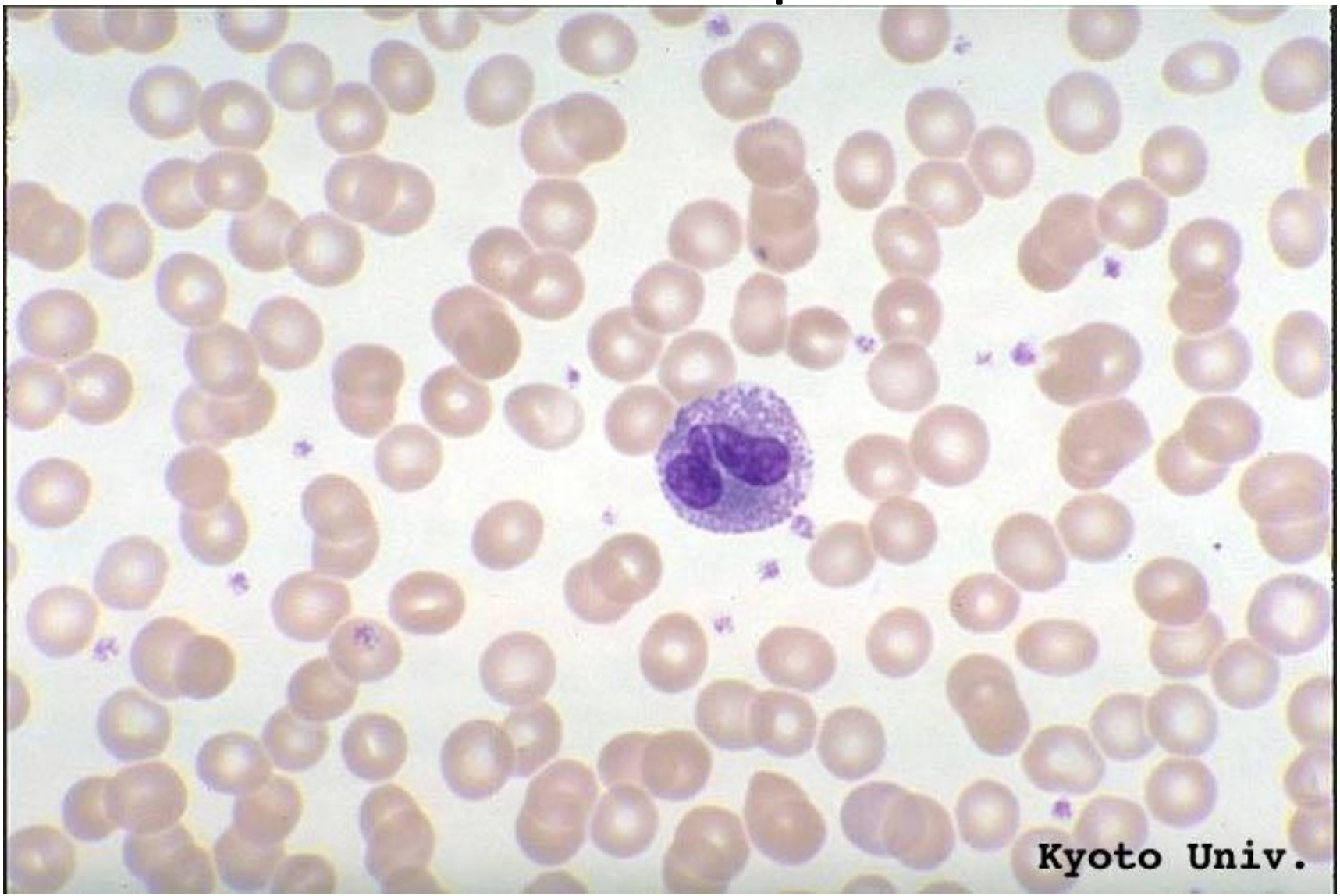
	SLIDE ID									
	H	G	C	B	D	I	F	A	E	J
PANCREAS										
LIVER										
KIDNEY										
LUNG										
TESTES										
NEUTROPHIL										
EOSINOPHIL										
BASOPHIL										
MONOCYTE										
MONOCYTE										

	NotApplic	Poor	Below Ave	Average	Above Average	Good
HISTOLOGY COMPETANCE						
HAEMATOLOGY COMPETANCE						

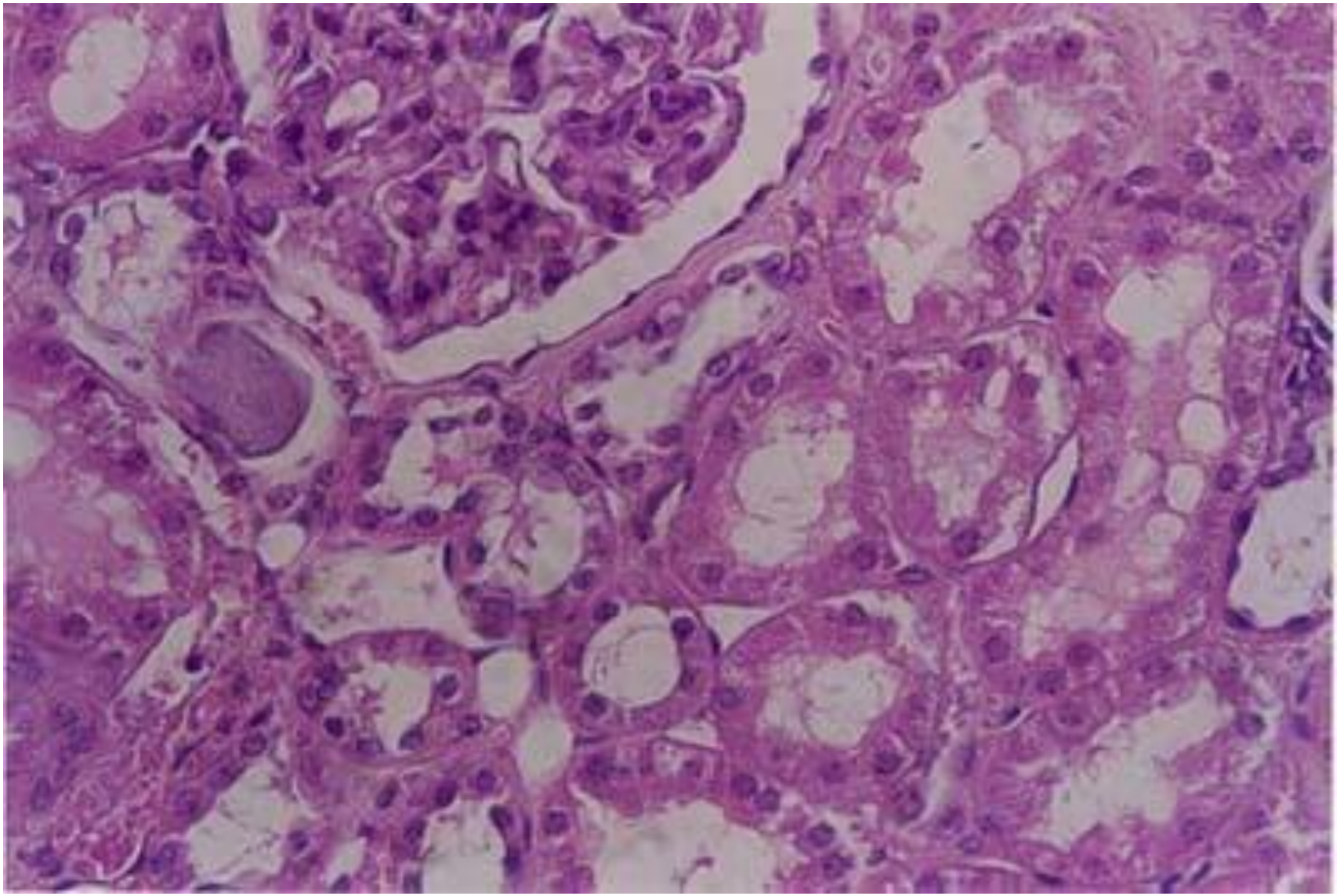
H: basophil



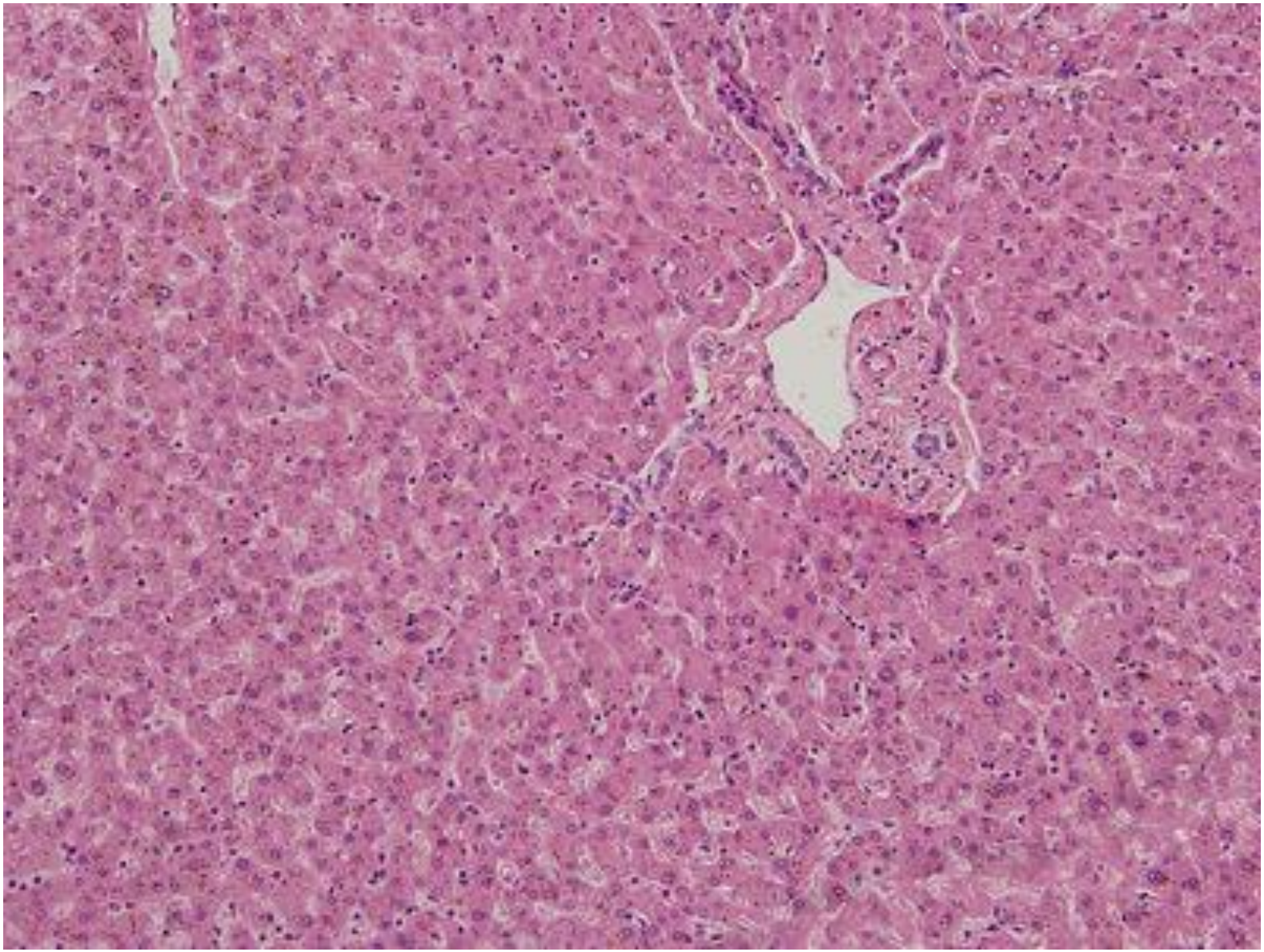
G : eosinophil



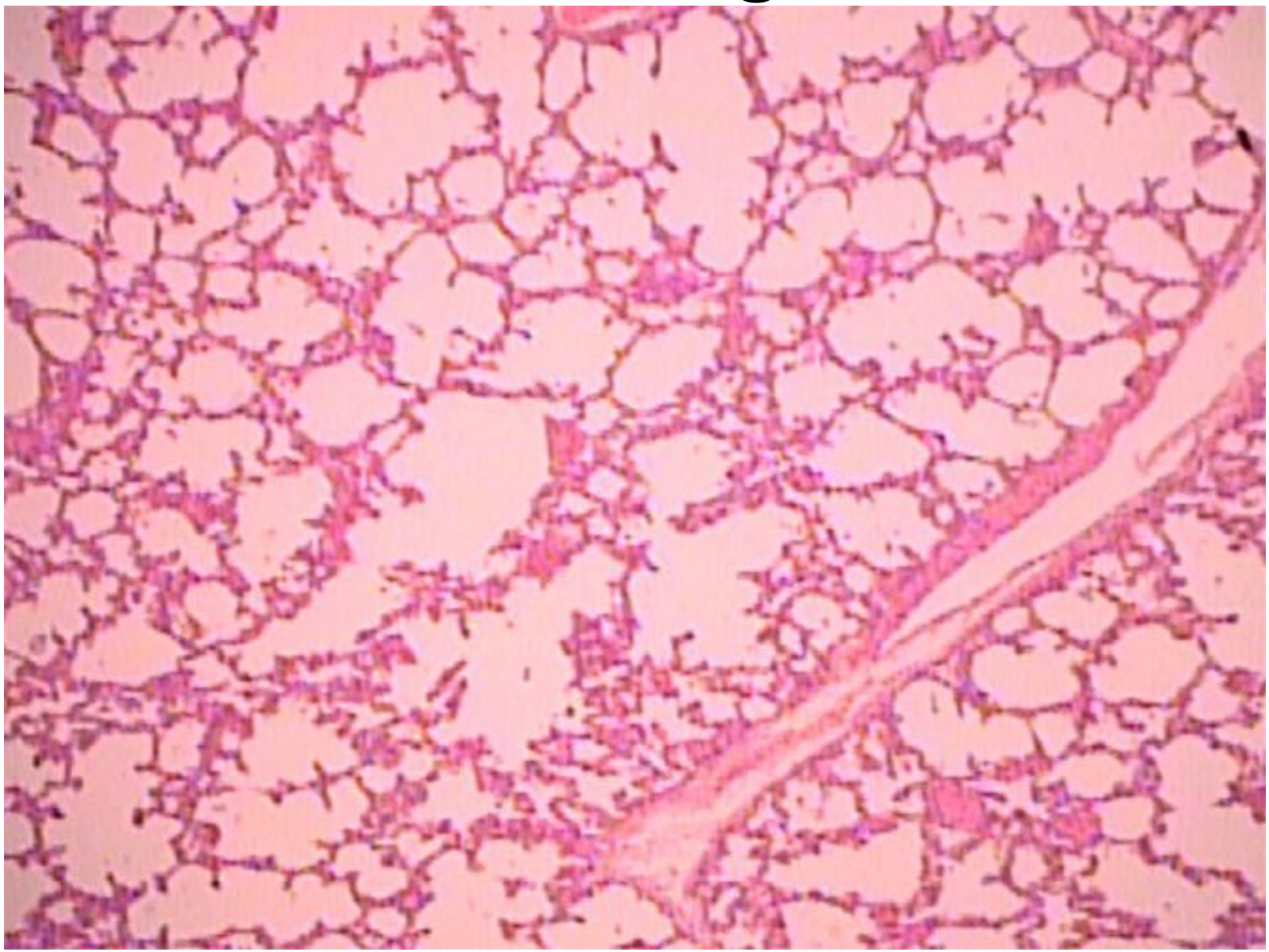
C : kidney



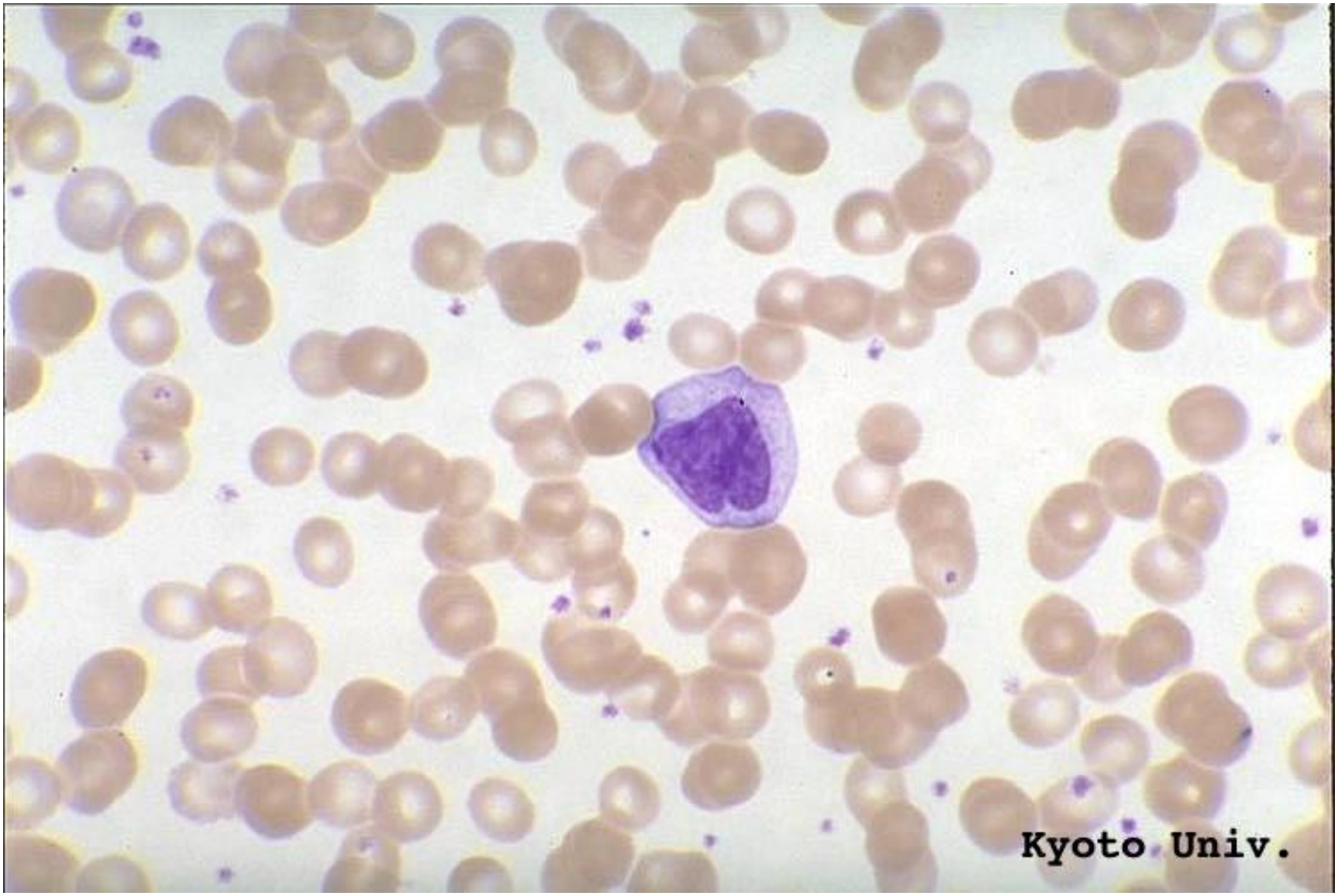
B : liver



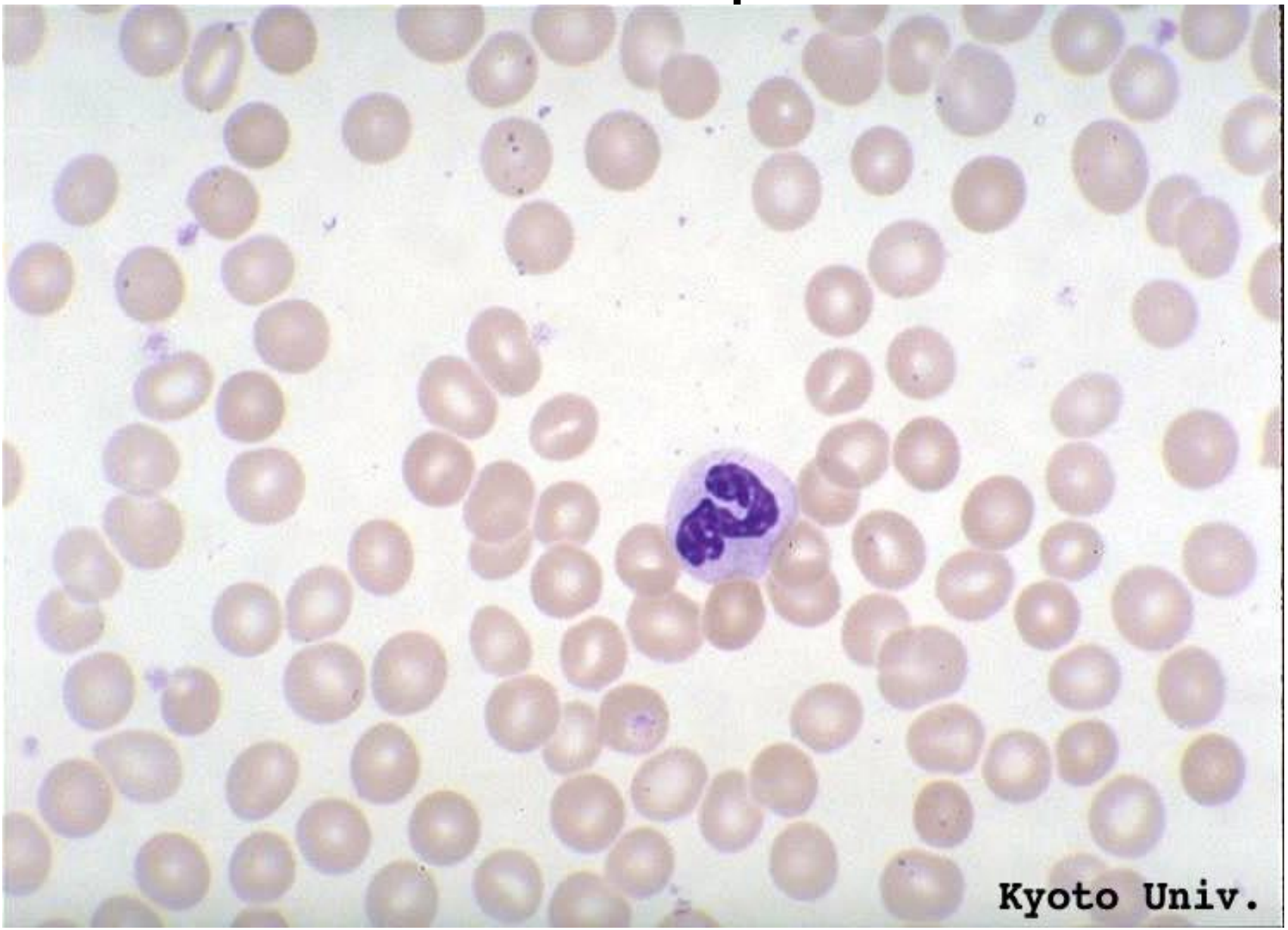
D : lung



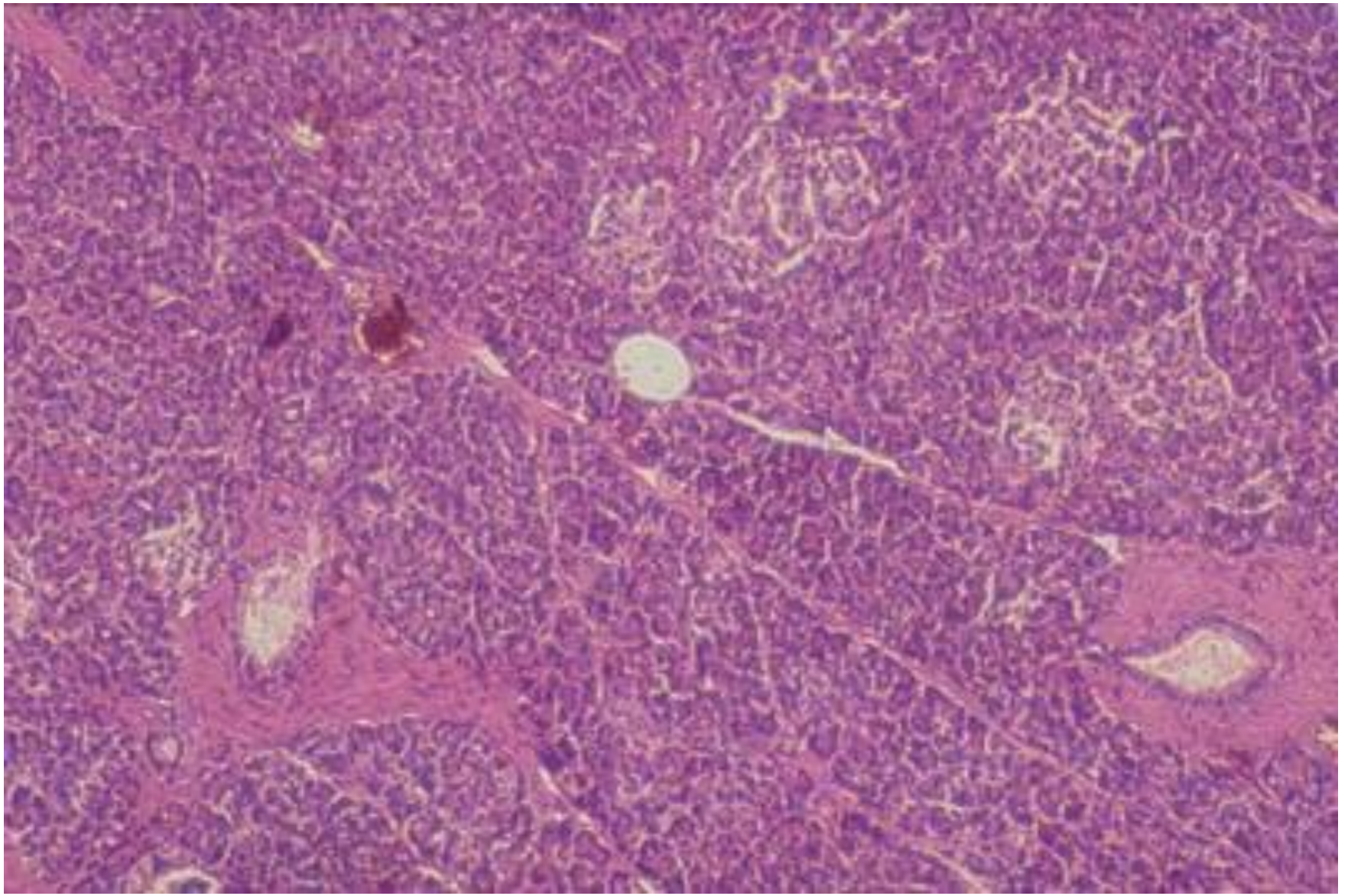
l : monocyte



F : neutrophil

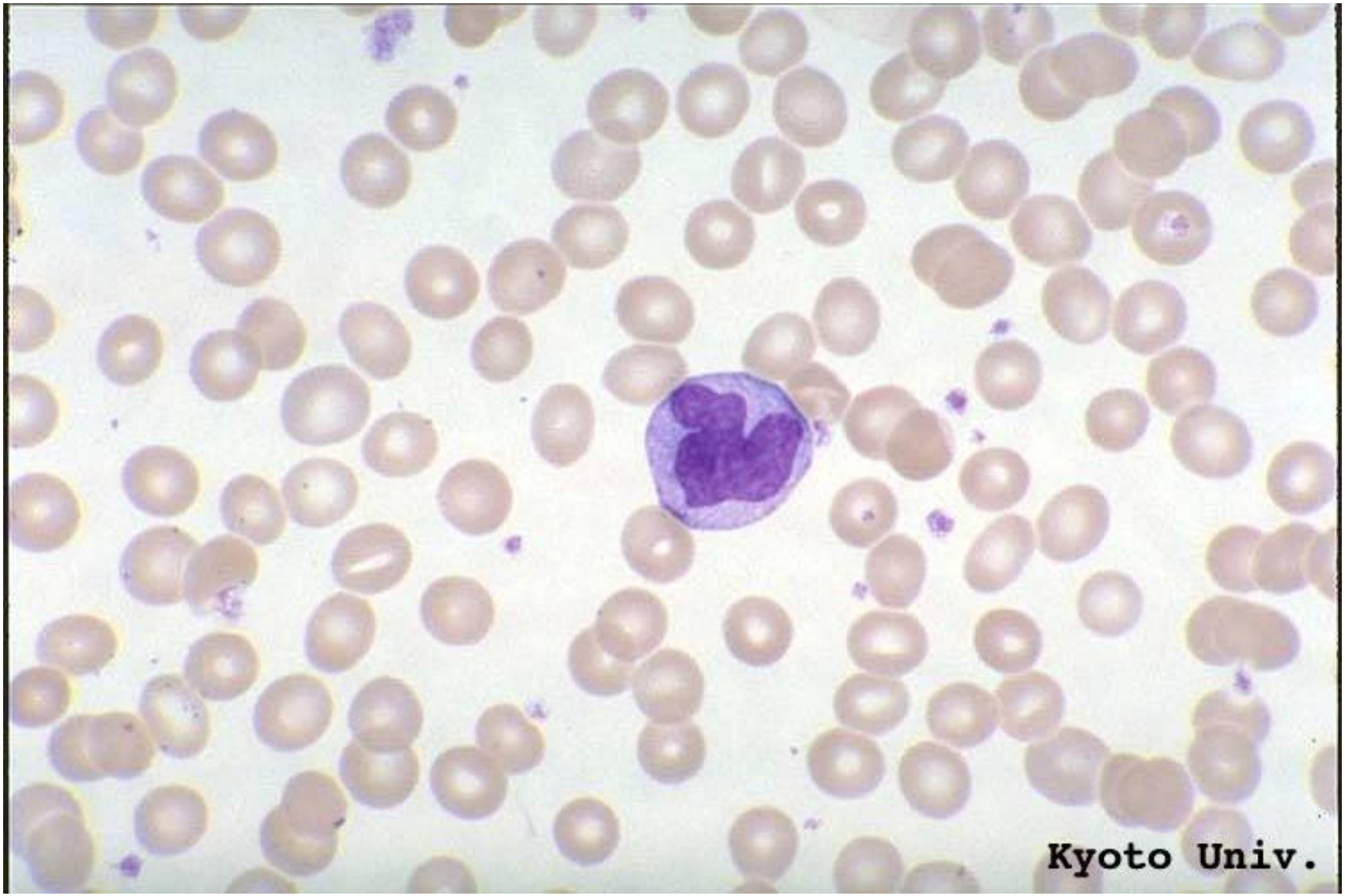


A : pancreas



E : testes

J : monocyte



Results from my Real Survey

	SLIDE ID										Probability of Right ID
	A	B	C	D	E	F	G	H	I	J	
PANCREAS : A	2	3	2	0	4	0	0	0	0	0	0.18
LIVER : B	5	6	1	0	0	0	0	0	0	0	0.46
KIDNEY : C	1	0	8	0	4	0	0	0	0	0	0.62
LUNG : D	0	1	0	11	0	0	0	0	0	0	1.00
TESTES : E	3	3	2	0	5	0	0	0	0	0	0.38
NEUTROPHIL : F	0	0	0	0	0	10	5	2	0	4	0.77
EOSINOPHIL : G	0	0	0	0	0	2	4	1	0	0	0.36
BASOPHIL : H	0	0	0	0	0	1	0	5	1	1	0.45
MONOCYTE : I	0	0	0	0	0	0	2	3	11	0	0.85
MONOCYTE : J	0	0	0	0	0	0	0	0	1	8	0.62
									Prob All	RIGHT =	0.0013

The Correct Identification rate was 57.4% overall.
The Correct 'Histo' Identification rate was 52%.
The Correct 'Haem' Identification rate was 62%.

The probability of a participant getting all identifications correct was 0.0013
The probability of a participant getting all Histo slides correct was 0.02
The probability of a participant getting all Haem slides correct was 0.07

HISTO
Pancreas and testes were the most poorly identified
Pancreas tended to be identified as liver
Testes tended to be identified as kidney or pancreas.

HAEM
Eosinophils were the most poorly identified
Eosinophils tended to be identified as neutrophils.
Monocytes (slide J) tended to be identified as neutrophils.

Overall the results reflected the skills mix in the audience which was predominantly 'haematological' staff and 'multidisciplinary' Core Lab staff.

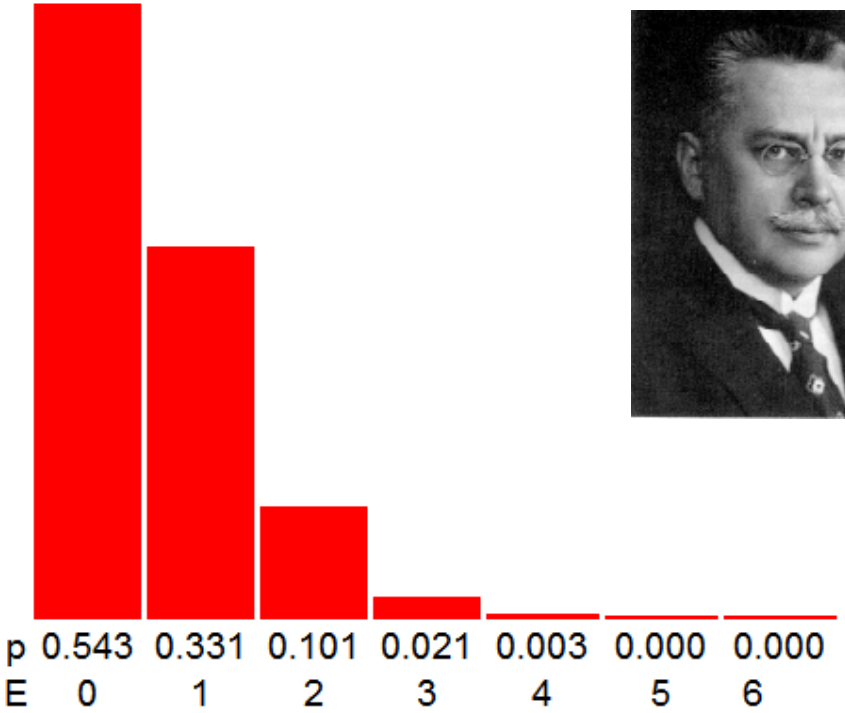
Weakness of my study – did not put in a slide of a tissue that was not on the list with the option on the survey – 'Tissue type not on the list of options'. If that had been there then I could have calculated all the requirements of Technical Note 17

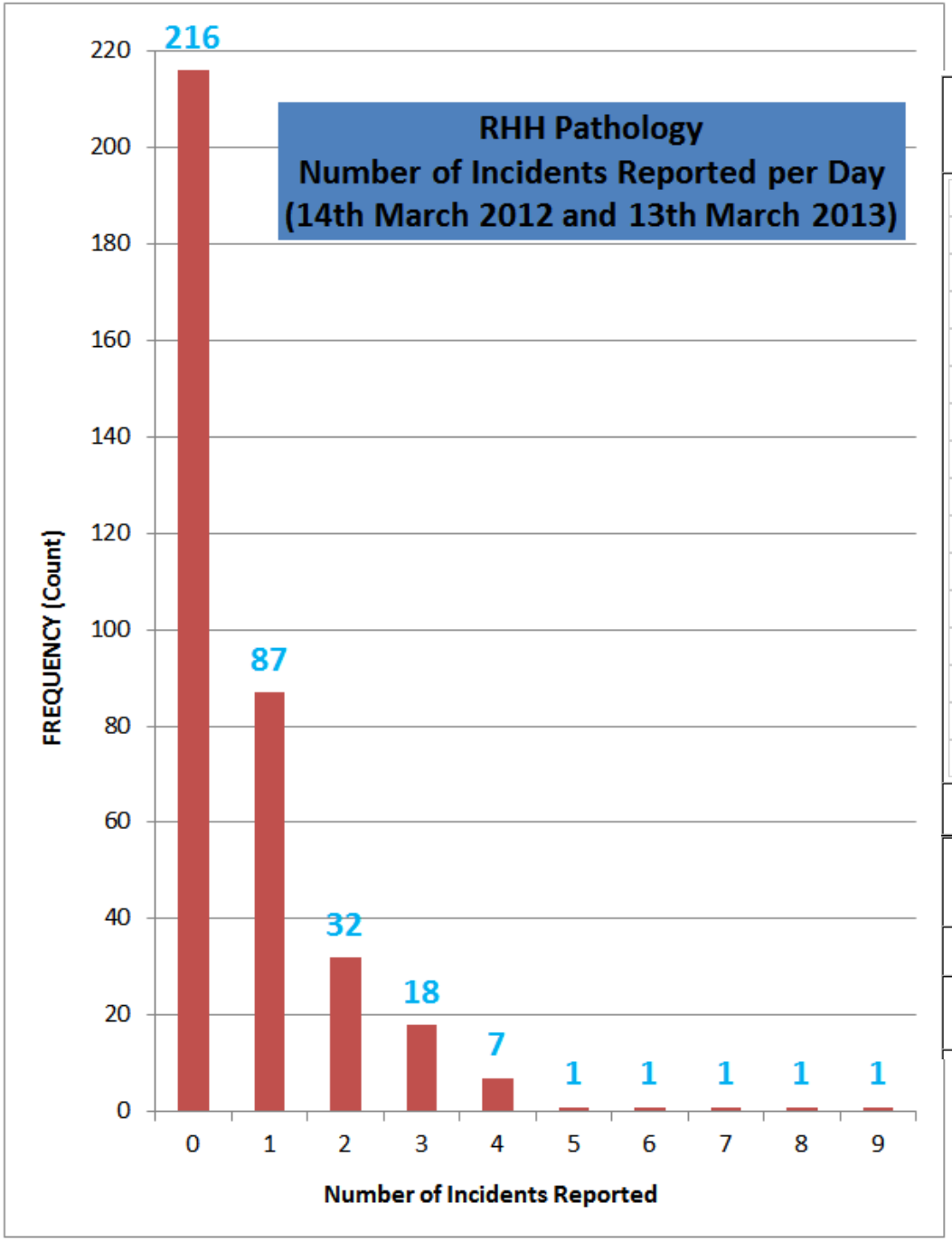
WHAT IS LIMITING OUR ABILITY TO COLLECT “REAL DATA” ?

THE STATISTICS OF RARE EVENTS IS THE INSURMOUNTABLE LIMIT.

The classic Poisson example is the data set of Ladislaus von Bortkiewicz (1898), for the chance of a Prussian cavalryman being killed by the kick of a horse. Ten army corps were observed over 20 years, giving a total of 200 observations of one corps for a one year period. The period or module of observation is thus one year. The total deaths from horse kicks were 122, and the average number of deaths per year per corps was thus 122/200 = 0.61…… Here, then, is the classic Poisson situation: a rare event, whose average rate is small, with observations made over many small intervals of time.

Simeon-Denis Poisson 1781-1840





k	Observed		Fitted Poisson	
	Frequency	Proportion	Probability	Expected Frequency
0	216	0.5918	0.61263	223.6086
1	87	0.2384	0.30019	109.5682
2	32	0.0877	0.07355	26.8442
3	18	0.0493	0.01201	4.3846
4	7	0.0192	0.00147	0.5371
5	1	0.0027	0.00014	0.0526
6	1	0.0027	0.00001	0.0043
7	1	0.0027	0	0.0003
8	1	0.0027	0	0
9	1	0.0027	0	0
10				
11				
12				
13				
14				
15				

Reset Calculate

mean of observed sample =	0.73
variance of observed sample =	1.5
mean and variance of fitted Poisson distribution =	0.49
observed vs expected frequencies proportion of linear covariance: r^2 =	0.98803

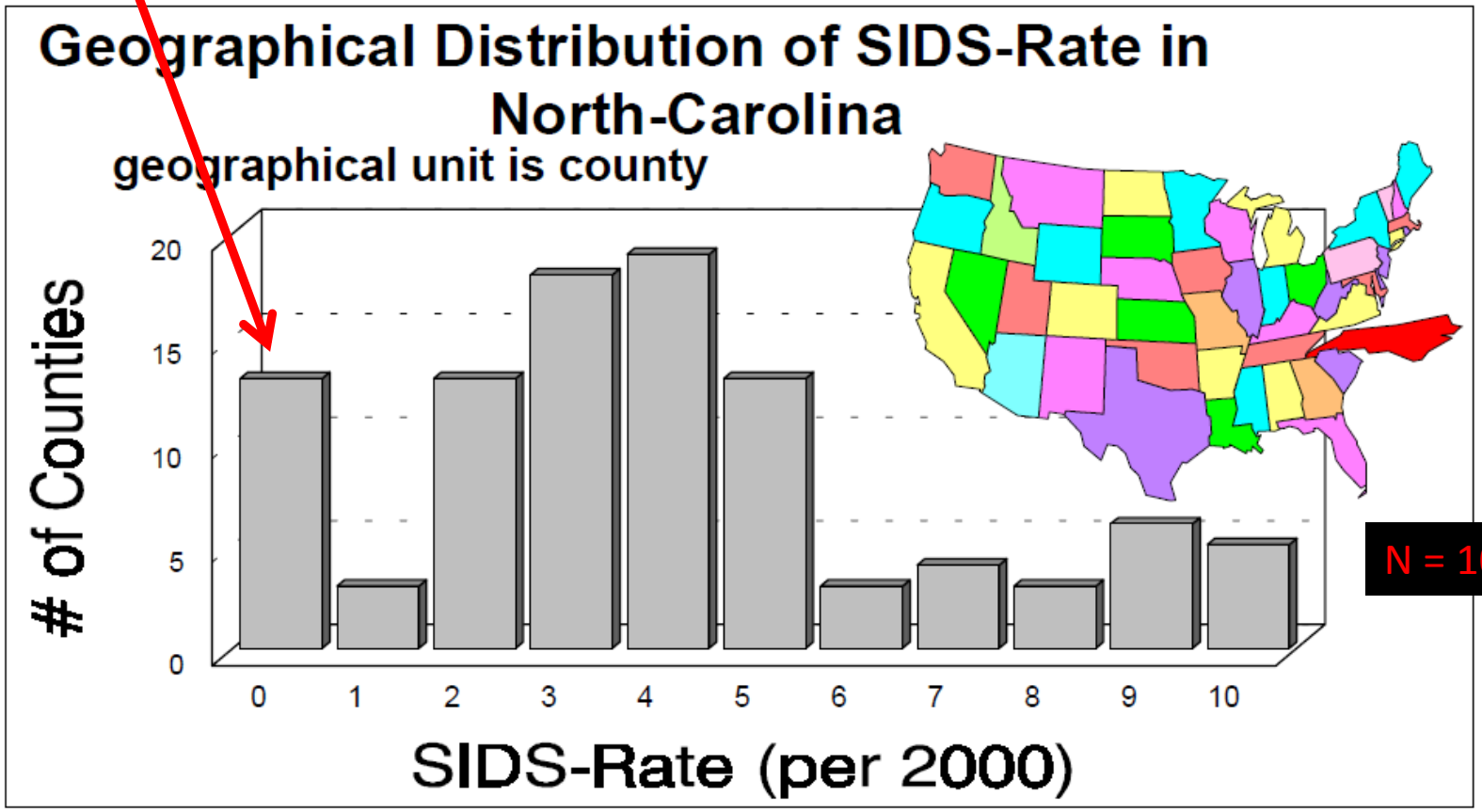
ANOTHER CANDIDATE STATISTICAL MODEL IS THE ZERO INFLATED POISSON (ZIP) DISTRIBUTION

Chapter 32

Zero-Inflated Count Models and their Applications in Public Health and Social Science

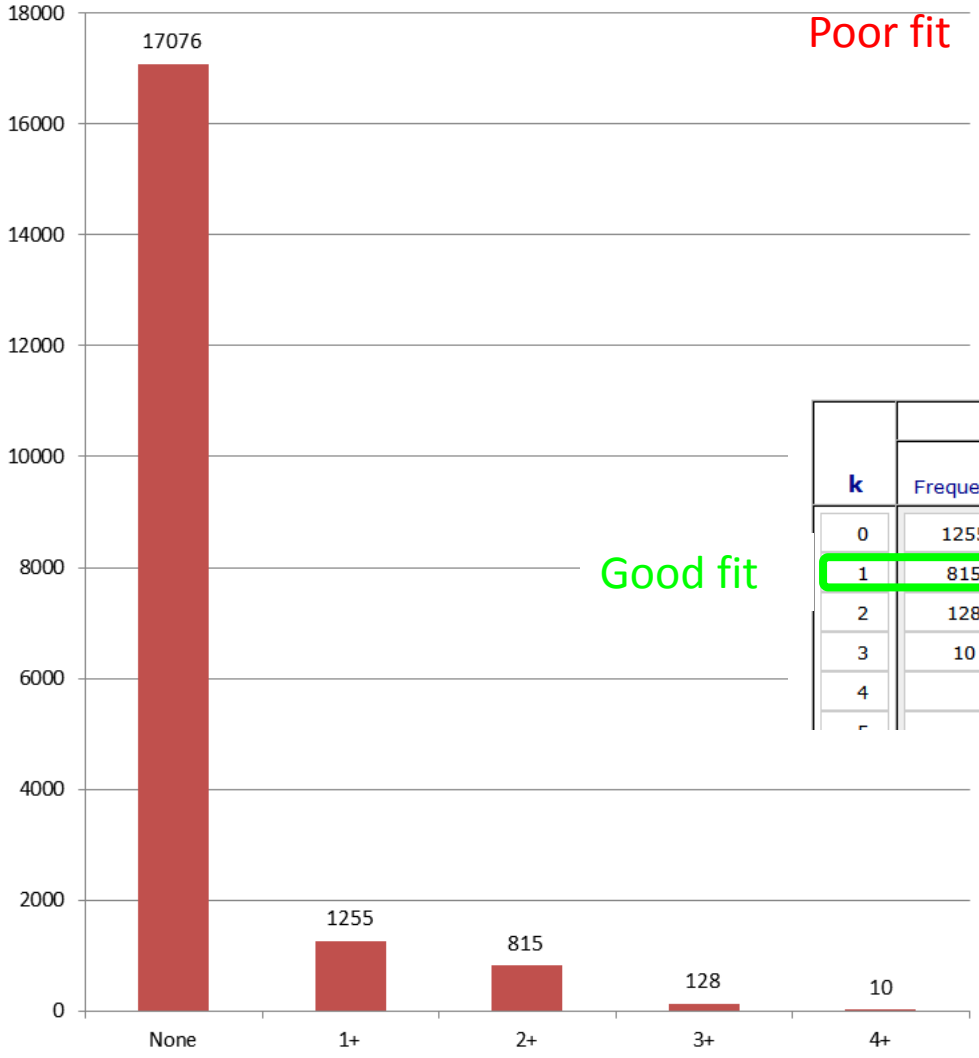
Dankmar Böhning, Ekkehart Dietz and Peter Schlattmann

Department of Epidemiology, Institute for Social Medicine, Free University Berlin



N = 100 counties

Figure 6: Geographical distribution of SIDS-rate in north-carolina



Poor fit

k	Observed Frequency	Observed Proportion	Fitted Poisson Probability	Fitted Poisson Expected Frequency
0	17076	0.8855	0.90484	17448.8848
1	1255	0.0651	0.09048	1744.8885
2	815	0.0423	0.00452	87.2444
3	128	0.0066	0.00015	2.9081
4	10	0.0005	0	0.0727
5				
6				
7				
8				
9				
10				

Good fit

k	Observed Frequency	Observed Proportion	Fitted Poisson Probability	Fitted Poisson Expected Frequency
0	1255	0.5684	0.56553	1248.6802
1	815	0.3691	0.32235	711.7477
2	128	0.058	0.09187	202.8481
3	10	0.0045	0.01746	38.5411
4				

:	0.17
variance of observed sample =	0.27
mean and variance of fitted Poisson distribution =	0.1
observed vs expected frequencies proportion of linear covariance: r^2 =	0.99683

CONCLUSION – OUR HAEMOLYSIS RATE PROFILE IS PROBABLY ‘ZERO INFLATED’

Whatever the model … all risks … positive and negative when summed together must equal ONE !

The real way ahead for TOTAL RISK ANALYSIS is to adopt the Bayesian Network approach …..

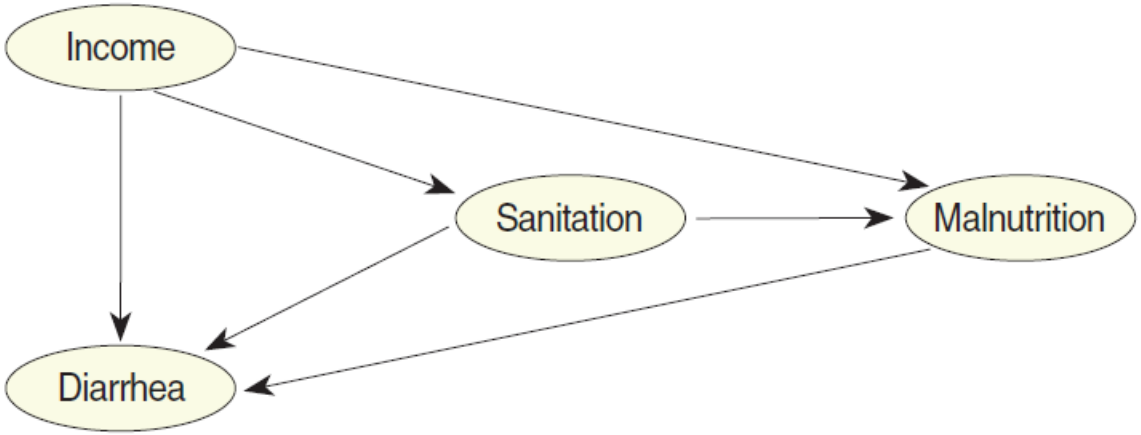
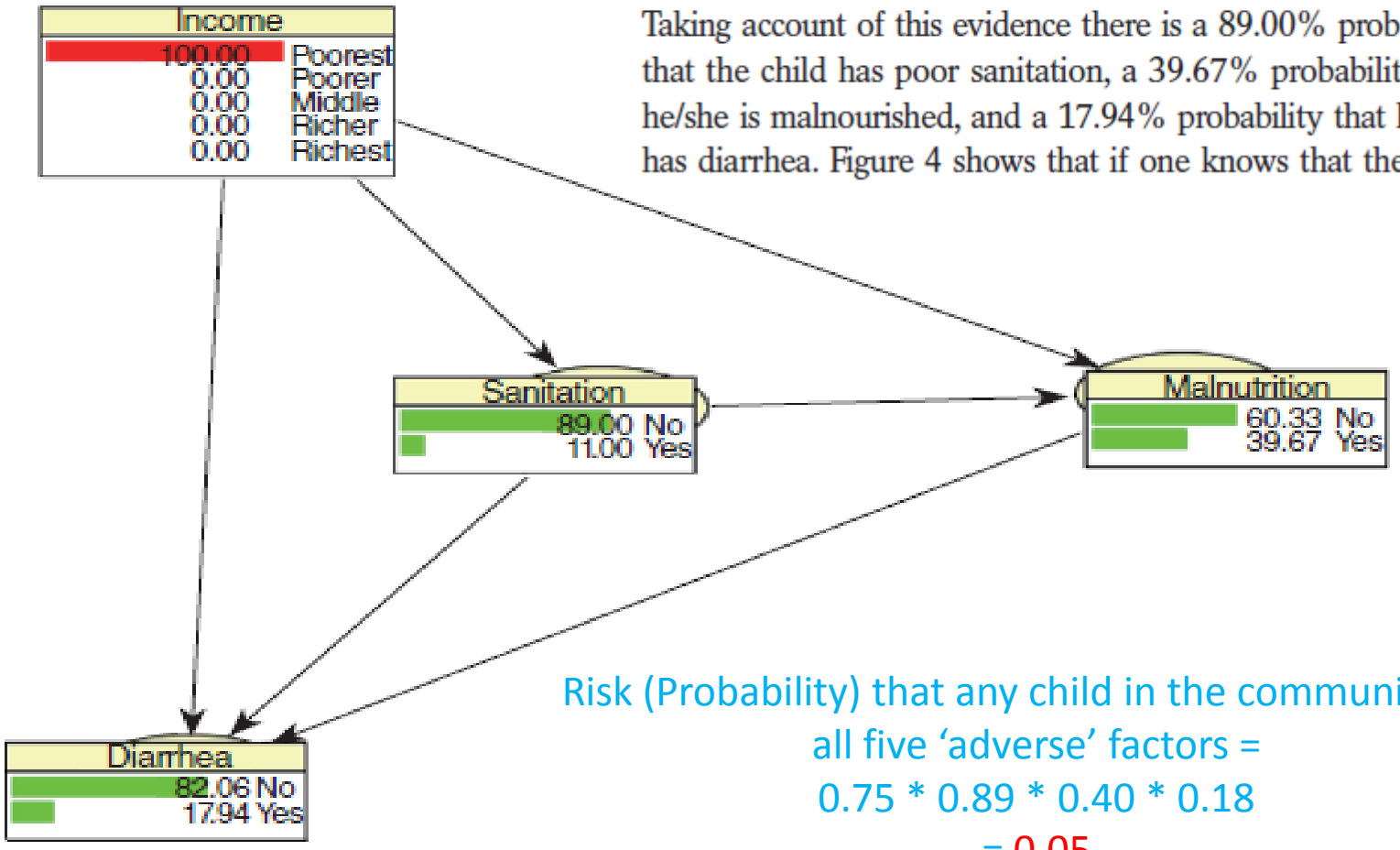


Figure 1. Bayesian network: a simplified conceptual hierarchical framework for diarrhea.

can also be expressed in terms of probabilities, e.g. that there is a 75% chance that the child's family falls in the poorest group). Taking account of this evidence there is a 89.00% probability that the child has poor sanitation, a 39.67% probability that he/she is malnourished, and a 17.94% probability that he/she has diarrhea. Figure 4 shows that if one knows that the child



Risk (Probability) that any child in the community has all five 'adverse' factors =
0.75 * 0.89 * 0.40 * 0.18
= 0.05

Figure 3. Frequency network showing posterior probabilities (%) when there is evidence that the child belongs to a family in the poorest quintile.

CONCLUSIONS

- The 'risk and harm' literature has so many different types of reports in it you can almost certainly find one to support any personal (subjective) opinion
- Nevertheless there are serious causes for concern
- Pre and Post analytical risk rates are higher than I anticipated (KIMMS)
- Risks in Quantitative analytical testing are very low
- Risks in Qualitative analytical testing are 'unknown' and untested in any meaningful study to date
- Risk data collection in any healthcare setting only makes sense if there is a model risk profile to test it against. (A HYPOTHESIS !) On first inspection the Poisson and Zero Inflated Poisson Distributions appear to be good starting candidates because there is a body of understanding of comparable processes in nature, epidemiology and production engineering.
- When related events are involved the a Bayesian Network is the way to go.
- Because adverse events are numerically rare the data collection needs to be automated wherever possible. The reliance upon 'self reporting' of adverse events will inevitably lead to under reporting of the 'numerator', unreliable estimation of the 'denominator' and the calculation of a distorted statistic.
- Once we have completed targeted data collection and analysis we can then embark on evidence based Quality Improvement.

Thank you for your attention