

melatonin suggests that it is well tolerated at the dose used. The present study, however, does not indicate whether it is necessary to take melatonin before and during the flight or only after it. Further research is needed on the dose response characteristics of melatonin to optimise its effect in alleviating jet lag.

Overall the results support the use of melatonin for alleviating jet lag and tiredness after long haul flights and indicate further investigations necessary to maximise the positive effects of melatonin.

We thank Air New Zealand for its help and support.

- 1 Arendt J, Marks V. Physiological changes underlying jet lag. *Br Med J* 1982;284:144-6.
- 2 Weaver RA. *The circadian system of man*. New York: Springer, 1979.
- 3 Nicholson AN, Pascoe PA, Spencer MB, Stone BM, Roehrs T, Roth T. Sleep after transmeridian flights. *Lancet* 1986;ii:1205-8.
- 4 Arendt J, Aldhous M, Marks V. Alleviation of jet lag by melatonin: preliminary results of a controlled double blind trial. *Br Med J* 1986;292:1170.
- 5 Arendt J, Aldhous M, English J, Marks V, Arendt JH. Some effects of jet-lag and their alleviation by melatonin. *Ergonomics* 1987;30:1379-93.
- 6 Arendt J, Marks V. Can melatonin alleviate jet lag? *Br Med J* 1983;287:426.
- 7 McNair DM, Lorr M, Droppelman LF. *Edits manual for the profile of mood states*. San Diego: EDITS, 1981.
- 8 SPSS Incorporated. *Users guide to statistical package for the social sciences—X*. New York: McGraw-Hill, 1983.
- 9 Anonymous. Jet lag and its pharmacology [Editorial]. *Lancet* 1986;ii:493-5.

(Accepted 2 December 1988)

## Observer variation in histopathological diagnosis and grading of cervical intraepithelial neoplasia

Sezgin M Ismail, Angela B Colclough, John S Dinnen, Douglas Eakins, David M D Evans, Ernest Gradwell, Jerry P O'Sullivan, Joan M Summerell, Robert G Newcombe

Department of  
Histopathology, University  
Hospital of Wales, Cardiff  
CF4 4XW

Sezgin M Ismail, MRCPATH,  
senior registrar in  
histopathology

Bolton General Hospital,  
Bolton, Lancashire BL4  
0JR

Angela B Colclough,  
MRCPATH, consultant  
histopathologist

County Hospital, Hereford  
John S Dinnen, MRCPATH,  
consultant histopathologist

Queen Elizabeth Hospital,  
King's Lynn, Norfolk  
Douglas Eakins, FRCPATH,  
consultant histopathologist

Llandough Hospital,  
Penarth, South Glamorgan  
David M D Evans, FRCPATH,  
consultant histopathologist

Whiston Hospital, Prescott,  
Merseyside L35 5DR  
Ernest Gradwell, MRCPATH,  
consultant histopathologist

St Richard's Hospital,  
Chichester PO19 4SE  
Jerry P O'Sullivan,  
FRCPATH, consultant  
histopathologist

Singleton Hospital,  
Swansea  
Joan M Summerell,  
FRCPATH, consultant  
histopathologist

Department of Medical  
Computing and Statistics,  
University of Wales  
College of Medicine,  
Cardiff  
Robert G Newcombe, PHD,  
lecturer

Correspondence to: Dr  
Ismail.

*Br Med J* 1989;298:707-10

### Abstract

To assess the variability among histopathologists in diagnosing and grading cervical intraepithelial neoplasia eight experienced histopathologists based at different hospitals examined the same set of 100 consecutive colposcopic cervical biopsy specimens and assigned them into one of six diagnostic categories. These were normal squamous epithelium, non-neoplastic squamous proliferations, cervical intraepithelial neoplasia grades I, II, and III, and other. The histopathologists were given currently accepted criteria for diagnosing and grading cervical intraepithelial neoplasia and asked to mark their degree of confidence about their decision on a visual linear analogue scale provided. The degree of agreement between the histopathologists was characterised by kappa statistics, which showed an overall poor agreement (unweighted kappa 0.358). Agreement between observers was excellent for invasive lesions, moderately good for cervical intraepithelial neoplasia grade III, and poor for cervical intraepithelial neoplasia grades I and II (unweighted kappa 0.832, 0.496, 0.172, and 0.175, respectively); the kappa value for all grades of cervical intraepithelial neoplasia taken together was 0.660. The most important source of disagreement lay in the distinction of reactive squamous proliferations from cervical intraepithelial neoplasia grade I. The histopathologists were confident in diagnosing cervical intraepithelial neoplasia grade III and invasive carcinoma (other) but not as confident in diagnosing cervical intraepithelial neoplasia grades I and II and glandular atypia (other).

Experienced histopathologists show considerable interobserver variability in grading cervical intraepithelial neoplasia and more importantly in distinguishing between reactive squamous proliferations and cervical intraepithelial neoplasia grade I. It is suggested that the three grade division of cervical intraepithelial neoplasia should be abandoned and a borderline category introduced that entails follow up without treatment.

### Introduction

Cervical intraepithelial neoplasia is the name given to a range of squamous epithelial abnormalities of the cervix uteri that are associated with an increased risk of subsequent invasive carcinoma of the cervix. In current histopathological practice lesions are graded

according to the degree of differentiation as cervical intraepithelial neoplasia grade I, II, or III. Cervical intraepithelial neoplasia grade I represents the best differentiated lesions, previously categorised as mild dysplasia, and grade III the poorly differentiated lesions corresponding in the old terminology to severe dysplasia or carcinoma *in situ*. Cervical intraepithelial neoplasia grade II is the intermediate category equivalent to moderate dysplasia. The morphological criteria used in the diagnosis and grading of lesions have been reviewed by Buckley *et al.*<sup>1</sup>

Implicit in the nomenclature of cervical intraepithelial neoplasia is the concept of tumour progression and progressive loss of differentiation with increasing malignancy, which fits in well with the current understanding of the biology of carcinogenesis. It is presumed that at least some cases of cervical intraepithelial neoplasia grade I will eventually progress, probably through the higher grades of cervical intraepithelial neoplasia, to invasive carcinoma. Generally, cervical intraepithelial neoplasia grade I will probably take longer to progress to invasive carcinoma than a grade III lesion,<sup>2</sup> but forecasting the probable rate of progression is impossible in individual cases. The size of the lesion is considered to be a more important prognostic indicator than its histological grading. Therefore, the Ninth Study Group of the Royal College of Obstetricians and Gynaecologists has recommended that all cervical intraepithelial neoplasia should be regarded as a continuum and that cervical intraepithelial neoplasia grade I should be treated as seriously as grade III lesions.<sup>3</sup> Hence the diagnosis of cervical intraepithelial neoplasia of any grade may have serious treatment implications for the patient.

This study was designed to test the ability of histopathologists to distinguish consistently between cervical intraepithelial neoplasia and the reactive proliferations of cervical squamous epithelium and to assess the variability of grading of cervical intraepithelial neoplasia among different histopathologists. A further study designed to test the degree of intra-observer variation is currently in progress.

### Materials and methods

A total of 100 consecutive colposcopic biopsy specimens of the cervix—excluding those with a computer code indicating invasive carcinoma—that were received by the histopathology department at the University Hospital of Wales between 12 November

and 10 December 1986 were examined by a panel of eight histopathologists; the specimens were received fixed in 10% formaldehyde solution and were processed routinely. The same set of slides was given to all the histopathologists, the slides having been "blinded" by independent observers, who retained the key until the end of the study. The histopathologists were based at different hospitals and were all actively working in reporting general histopathological material, with a mean of 19 years' experience in histopathology (range 7-35 years).

#### DIAGNOSTIC CATEGORIES

The histopathologists (observers) allocated the biopsy specimens to one of six diagnostic categories. These were cervical intraepithelial neoplasia grades I, II, and III; normal squamous epithelium; and non-neoplastic reactive squamous proliferations—that is, squamous metaplasia, basal cell hyperplasia, inflammatory changes, and changes due to human papillomavirus without cervical intraepithelial neoplasia—and other, which included any abnormality of the cervical epithelium not covered by the preceding five categories.

The histopathologists were given definitions of these six diagnostic categories as well as a brief summary of the criteria generally used in the grading of cervical intraepithelial neoplasia. In addition, they were referred to the paper by Buckley *et al*<sup>1</sup> in which these criteria are described in detail. Every time they assigned a biopsy specimen to a diagnostic category the observers were asked to mark their degree of confidence on a visual linear analogue scale supplied as a matter of course. When they opted for other they were asked to specify the abnormality.

When the results were received it was found that observers opting for other specified either invasive carcinoma or glandular atypia of varying degrees of severity. Other was therefore subdivided into two categories, invasive carcinoma and glandular atypia, yielding altogether seven diagnostic categories for statistical analysis.

At the end of the study the original reports on the two cases regarded by most observers as invasive carcinoma were checked. Both had been originally reported as invasive carcinoma but had been miscoded.

#### STATISTICAL METHODS

The degree of agreement between arbitrary pairs of observers was characterised by kappa statistics.<sup>4,5</sup> Kappa is an index of interobserver agreement that has been corrected for chance and therefore indicates the degree of interobserver agreement over and above that which would be expected by chance alone. For most purposes kappa values greater than 0.75 may be taken to represent excellent agreement beyond chance, values below 0.40 poor agreement beyond chance, and values between 0.40 and 0.75 fair to good agreement beyond

chance.<sup>6</sup> Kappa values appropriate for multiple ratings for each observer were calculated for each of the seven diagnostic categories and for the overall agreement (see table I).

In addition a weighted kappa<sup>7</sup> was calculated (see table II). This is based on the idea that in any ordered scale composed of categories representing increasing severity of abnormality some possible disagreements are more serious than others. For example, in this study a paired observation with invasive carcinoma on the one hand and normal squamous epithelium on the other would represent a far more serious disagreement than one concerning neighbouring categories such as cervical intraepithelial neoplasia grades II and III. Whereas the unweighted kappa attributes equal importance to all disagreements, the weighted kappa takes into account the degree to which disagreements concern neighbouring categories.

The Kruskal-Wallis test was used to analyse the variation in the propensity of individual observers to assign optimistic versus pessimistic diagnostic grades (see table III) and to analyse the variation in the degree of confidence with which a biopsy specimen was allocated to each of the seven diagnostic categories (see table IV).

#### Results

Eight observers reading 100 specimens produced a total of 800 results, which are summarised in the

TABLE II—Unweighted and weighted kappa statistics for classification of 100 cervical biopsy specimens

	No of pairs of readings	Observed agreement	Chance expected agreement	Kappa
All data	5600	0.517	0.248	0.358
Excluding pairs of readings incorporating glandular atypia	5488	0.528	0.252	0.369
Using Fleiss-Cohen weights	5488	0.967	0.848	0.784

TABLE III—Variation among observers in categorising 100 cervical biopsy specimens. Values are numbers of observations

Observer No	Normal	Reactive	Cervical intraepithelial neoplasia			Total (I-III)	Invasive carcinoma	Glandular atypia*
			I	II	III			
1	10	41	12	19	16	47	1	1
2	11	40	15	17	9	41	2	6
3	8	46	10	14	20	44	2	
4	18	35	7	7	31	45	2	
5	8	37	15	20	19	54	1	
6	10	52	10	19	7	36	1	1
7	2	29	12	23	32	67	2	
8	14	40	11	15	18	44	2	

\*Excluded before application of the Kruskal-Wallis test.

TABLE I—Agreement between eight observers in categorising 100 cervical biopsy specimens

	No of individual observations	No of paired observations							Kappa for cervical intraepithelial neoplasia I-III	
		Cervical intraepithelial neoplasia		Cervical intraepithelial neoplasia			Total			
		Normal	Reactive	I	II	III				
Normal	81	250	289	9	8	1	10	567	0.378	
Reactive	320	289	1474	258	157	26	36	2240	0.430	
Cervical intraepithelial neoplasia:										
I	92	9	258	172	143	57	5	644	0.172	
II	134	8	157	143	294	333	3	938	0.175	
III	152	1	26	57	333	630	15	1064	0.496	
Invasive carcinoma	13					15	76	91	0.832	
Glandular atypia	8	10	36	5	3	2		56	-0.010	
Total	800	567	2240	644	938	1064	91	5600	0.358	0.517

TABLE IV—Diagnostic confidence measured by score out of 100 on visual analogue scale according to diagnosis

Diagnosis	No of observations (n=794)*	Score on visual linear analogue scale	
		Median	Range
Normal	80	86	13-97
Reactive	316	87	4-99
Cervical intraepithelial neoplasia:			
I	91	70	26-95
II	134	77	15-96
III	152	90	5-99
Invasive carcinoma	13	94	74-100
Glandular atypia	8	50	38-91

\*Diagnostic confidence was not scored for six of the 800 observations.

first column of table I. The categories from normal squamous epithelium to invasive carcinoma were judged to form an ordered scale representing increasing severity of abnormality. In eight instances the observer felt unable to assign the biopsy specimen to one of these six categories; these eight readings were classified as glandular atypia and do not form part of the ordinal scale.

To characterise the degree of agreement between any pairs of observers a symmetrical agreement matrix<sup>5</sup> was formed. This had a total of 5600 (100×8×7) possible pairs. Table I shows, for instance, that 81 of the individual assignments were categorised as normal squamous epithelium; each of these may be contrasted with each of the seven readings by other observers on the same specimen, yielding a total of 567 (81×7) comparisons. In 250 instances the other observer also classified the specimen as normal squamous epithelium, in 289 instances as reactive squamous proliferation, in nine instances as cervical intraepithelial neoplasia grade I, in eight as cervical intraepithelial neoplasia grade II, in one as cervical intraepithelial neoplasia grade I, and in 10 as glandular atypia.

The penultimate column of table I gives kappa values for each category and for the full seven category classification. A total of 2896 out of 5600 (51.7%) paired readings lay on the main diagonal representing exact agreement between the two assessors. We would expect, however, a certain degree of agreement by chance, even without any true consensus; this was estimated as 24.8% from the marginal frequency distribution. The multirater kappa statistic representing the degree of agreement beyond chance expectation was 35.8% (0.517-0.248)/(1-0.248), suggesting poor overall agreement.

Kappa statistics for individual categories were obtained similarly. The degree of agreement obviously varied widely among the categories. The degree of agreement for the few instances in which invasive carcinoma was identified was extremely high. The degree of agreement for the categories cervical intraepithelial neoplasia grade III, reactive squamous proliferations, and normal squamous epithelium was mediocre, whereas the degree of agreement for cervical intraepithelial neoplasia grades I and II was extremely poor, suggesting that these categories were not reliably discriminated from each other or their neighbours. The kappa statistic for all grades of cervical intraepithelial neoplasia taken together was 0.660. This is a considerable improvement on the kappa values obtained for each of the individual grades, suggesting that abolishing the three grades of cervical intraepithelial neoplasia would improve the degree of interobserver agreement in diagnosing the condition.

Table II shows the effect of assigning different penalty weights to different degrees of disagreement. Most of the disagreements seen in table I are by one or at most two categories: invasive carcinoma was never confused with normal or reactive squamous epithelium. A weighted kappa, with Fleiss-Cohen weights,

was constructed, excluding glandular atypia. This suggested closer agreement between observers than would be represented by the unweighted kappa, which disregards the hierarchical order.

Table III shows the variation among the eight observers in categorising the cervical biopsy specimens. The Kruskal-Wallis test was used to analyse these data after excluding the eight observations categorised as glandular atypia. When this was done the median grade was cervical intraepithelial neoplasia grade II for observer No 7, cervical intraepithelial neoplasia grade I for observer No 5, and reactive squamous epithelium for all the other observers. The Kruskal-Wallis test indicated highly significant variation among the eight observers ( $\chi^2=27.2$ ,  $p<0.001$ ), observer No 6 being significantly more optimistic and observer 7 significantly more pessimistic than the panel as a whole. The differences among observers are emphasised in the column showing the total number of specimens categorised as cervical intraepithelial neoplasia (grades I-III): observer No 7 had 67 such diagnoses and observer No 6 only 36.

Table IV summarises the confidence with which the observers assigned the biopsy specimens into the different diagnostic categories. The Kruskal-Wallis test indicated highly significant differences ( $\chi^2=105.4$ ,  $p<0.001$ ) in the degree of confidence associated with the seven diagnostic categories. Confidence was poorest for glandular atypia but was also relatively weak for cervical intraepithelial neoplasia grades I and II. In contrast invasive carcinoma and cervical intraepithelial neoplasia grade III were diagnosed with a high degree of confidence.

## Discussion

This study has shown poor interobserver agreement in the histopathological diagnosis and grading of cervical intraepithelial neoplasia. Agreement between observers was excellent for invasive lesions but poor in cervical intraepithelial neoplasia grades I and II and mediocre in grade III lesions. More important from the clinical point of view is the fact that a panel of experienced histopathologists had considerable difficulty in distinguishing cervical intraepithelial neoplasia grade I from reactive proliferations of the cervical squamous epithelium. Table I shows that, although there were 172 occasions when two observers agreed on a diagnosis of cervical intraepithelial neoplasia grade I, there were 258 occasions when one observer considered a lesion to be cervical intraepithelial neoplasia grade I and the other interpreted it as a non-neoplastic reactive proliferation. The difficulties experienced by the observers in distinguishing cervical intraepithelial neoplasia grade I from reactive squamous proliferations suggest that the currently recommended policy of treating cervical intraepithelial neoplasia grade I as seriously as grade III lesions may in some cases lead to overtreatment.

Similar observations were made by Bellina *et al*, who circulated 26 colposcopic cervical biopsy specimens among four "competent and respected histopathologists" from different institutions, of whom one was a general pathologist and three were gynaecological pathologists.<sup>8</sup> Comparisons between pathologists resulted in agreement only half of the time. Although most of the disagreements were within one grade of each other, a small proportion of them concerned normal squamous epithelium on the one hand and cervical intraepithelial neoplasia grade III on the other. Cocker *et al* studied variation among observers in the diagnosis of cervical squamous epithelial abnormalities in 28 selected cases in which the original diagnosis ranged from squamous cell carcinoma to squamous metaplasia and reserve cell hyperplasia.<sup>9</sup> Serious in-

consistencies in diagnosis were shown among three histopathologists with special experience of gynaecological pathology who had worked together in the same institution for at least two years.

The poor interobserver agreement in our study may simply reflect shortcomings in the diagnostic ability or training of the members of the panel. The histopathologists who took part in this study, however, were all active in routine diagnostic histopathology, had worked in different hospitals, and had considerable experience of histopathology. Participation in a lengthy study of this type necessitates considerable interest in self assessment and insight into the problems that prompted the study. The observers' awareness of the problems associated with the diagnosis and grading of cervical intraepithelial neoplasia is emphasised by the fact that invasive carcinoma was diagnosed with a high degree of confidence and excellent interobserver agreement, whereas cervical intraepithelial neoplasia grade I was diagnosed with a low degree of confidence and correspondingly poor interobserver agreement. Although a significant degree of systematic observer bias was seen, the diversity in the histopathological background of the members of the panel and their evident high motivation and awareness of problems combined with previous observations<sup>8,9</sup> suggest that inadequacy of the observers alone is an improbable explanation for our findings.

An alternative and more likely explanation is that the poor agreement among observers is largely caused by shortcomings in the morphological criteria used to diagnose cervical intraepithelial neoplasia and in the current grading system. Cervical intraepithelial neoplasia has been described as a continuum of abnormality; its subdivision into three grades creates two internal boundaries, and our results show that the location of these boundaries is highly subjective. Morphological criteria currently used in the diagnosis of cervical intraepithelial neoplasia probably identify a heterogeneous group of abnormalities rather than distinct histological stages in the development of cervical carcinoma. Morphological changes resembling cervical intraepithelial neoplasia can be induced by a non-carcinogenic chemical such as podophyllum resin; these changes regress when the chemical is withdrawn.<sup>10</sup> Nasiell *et al* found a rate of regression of 65% in mild dysplasia, 54% in moderate dysplasia, and 43% in severe dysplasia, the rate of progression being 23%, 30%, and 38% respectively.<sup>11,12</sup> More recently Robertson *et al* in a survey of 1781 women with mild dyskaryosis found a reversion to normal in 46% of the patients.<sup>13</sup> Therefore, additional methods are needed to improve the accuracy of diagnosis of the preinvasive stages of cervical carcinoma and to define the prognosis of any morphological abnormalities identified.

Several possible ways of improving our understanding of the natural course of the disease are currently being explored. Many varieties of human papillomavirus have been identified, and their role in the progression of cervical neoplasia is being studied; some subtypes of human papillomavirus infection may be associated with an increased risk of progression to invasive carcinoma. An American study on the prognostic importance of measurements of nuclear DNA showed that 91% of diploid or polyploid lesions regress and 9% persist.<sup>14</sup> In contrast, 81% of aneuploid lesions in this series persisted, 12% progressed to invasive carcinoma, and 7% showed normal results on follow up. The best morphological indicator of aneuploidy was the presence of atypical mitotic figures. It may well be useful to examine more sections from colposcopic cervical biopsy specimens in a search for

abnormal mitoses and to combine morphological assessment of low grade cervical intraepithelial neoplasia with nuclear DNA measurements.

Our results conclusively show that the current system of grading cervical intraepithelial neoplasia is poorly reproducible between observers and that there is considerable overlap between non-neoplastic squamous proliferations and the lower grades of cervical intraepithelial neoplasia. Accordingly, we suggest that an inaccurate poorly reproducible grading system, which in any case is not essential to current clinical management, should be altogether abandoned, that the three subdivisions of cervical intraepithelial neoplasia should be abolished, and that pathologists should be prepared to diagnose and gynaecologists to accept a borderline category.

The cervical cytology screening programme is detecting increasing numbers of women with minor cytological abnormalities who are subsequently referred for colposcopic biopsy with resulting increasing pressure on colposcopy and histopathology services. A borderline category of cervical intraepithelial neoplasia is in practice already being used by some histopathologists, and we believe that the formal introduction and widespread use of such a category with its clinical implications of follow up without treatment will enable us to elucidate the natural course of these minor abnormalities. Hence ultimately we may be able to relieve the increasing pressure on scarce health service resources by identifying the instances in which treatment is unnecessary.

We thank Professor E D Williams for his interest in this work and his helpful criticism of the manuscript.

- 1 Buckley CH, Butler EB, Fox H. Cervical intraepithelial neoplasia. *J Clin Pathol* 1982;35:1-13.
- 2 Barron BA, Richart RM. A statistical model of the natural history of cervical carcinoma based on a prospective study of 557 cases. *JNCI* 1968;41:1343-53.
- 3 Jordan JA, Sharp F, Singer A, eds. *Proceedings of the ninth study group of the Royal College of Obstetricians and Gynaecologists*. London: Royal College of Obstetricians and Gynaecologists, 1982:299-300.
- 4 Fleiss JL. *Statistical methods for rates and proportions*. 2nd ed. New York: Wiley, 1981:212-34.
- 5 Beck JS. Observer variability in reporting of breast lesions. *J Clin Pathol* 1985;38:1358-65.
- 6 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-74.
- 7 Fleiss JL, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educational and Psychological Measurement* 1973;33:613-9.
- 8 Bellina JH, Dunlap WP, Riopelle MA. Reliability of histopathologic diagnosis of cervical intraepithelial neoplasia. *South Med J* 1982;75:6-8.
- 9 Cocker J, Fox H, Langley FA. Consistency in the histological diagnosis of epithelial abnormalities of the cervix uteri. *J Clin Pathol* 1968;21:67-70.
- 10 Christopherson WM. Concepts of genesis and development in early cervical neoplasia. *Obstet Gynecol Surv* 1969;24:842-50.
- 11 Nasiell K, Nasiell M, Vaclavinkova V. Follow-up studies of cytologically detected precancerous lesions (dysplasia) of the uterine cervix. In: Boström H, *et al*, eds. *Skandia international symposia: health control in detection of cancer*. Stockholm: Almqvist and Wiksell International, 1976:244-52.
- 12 Nasiell K, Nasiell M, Vaclavinkova V. Behavior of moderate cervical dysplasia during long-term follow-up. *Obstet Gynecol* 1983;61:609-14.
- 13 Robertson JH, Woodend BE, Crozier EH, Hutchinson J. Risk of cervical cancer associated with mild dyskaryosis. *Br Med J* 1988;297:18-21.
- 14 Fu YS, Reagan JW, Richart RM. Definition of precursors. *Gynecol Oncol* 1981;12:S220-3.

(Accepted 16 January 1989)

## Correction

### Waldenström's macroglobulinaemia: three cases in shoe repairers

An editorial error occurred in the authors' names of this paper by Dr Lorna M Williamson and others (25 February, p 498). The third author is Jonathan R Worters and not Jonathan R Waters as published.