8 Vokes EE, Weichselbaum RR. Concomitant chemo-radiotherapy: rationale and clinical experience in patients with solid tumours. *J Clin Oncol* 1990;8:911-34.
9 Gupta NL, Pointon RCS, Wilkinson PM. A randomised clinical trial to contrast radiotherapy with radiotherapy and methotrexate given synchronously in head and neck cancer. *Clin Radiol* 1987;38:575-81.
10 Schaake-Koning C, van den Bogaert W, Dalesio O, Festen J, Hoogenhout J, van Houtte P, et al. Effects of concomitant cisplatin and radiotherapy on inoperable non-small-cell lung cancer. *N Engl J Med* 1992;326:524-30.
11 Furuse K, Fukuoka M, Kawahara M, Nishikawa H, Takada Y, Kudoh S, et al. Mitomycin, vindesine and cisplatin in unresectable stage III non-small-cell lung cancer. *J Clin Oncol* 1999;17:2692-9.
12 Overgaard J, Sand Hansen H, Sapru W, Overgaard M, Grau C, Specht L, et al. The DAHANCA 6 & 7 trial of 5 vs. 6 fractions per week of conventional radiotherapy of squamous cell carcinoma of the head and neck: a randomised study with 1485 patients. *Radiother Oncol* 2001;58 (suppl 1):S40.
13 Saunders M, Dische S, Barrett A, Harvey A, Gibson D, Parmar M, et al. Continuous hyperfractionated accelerated radiotherapy (CHART) vs. conventional radiotherapy in non-small-cell lung cancer: a randomised multicentre trial. *Lancet* 1997;350:161-5.

HPV testing for clarifying borderline cervical smear results

Recent conflicting results highlight the dilemmas of progress

Two large studies in the United States have provided strong evidence that testing for human papillomavirus DNA is a useful tool for managing women with borderline results on cervical smear tests.[1 2] In response to those results and subsequent cost analyses, integrated managed care organisations, including Kaiser Permanente of California, have recently implemented human papillomavirus testing for managing women with atypical squamous cells of undetermined significance (ASCUS) on cervical smear tests. In the past year many health insurance plans have done the same. Consequently, a large majority of insured women in the United States now have cover for human papillomavirus testing in response to atypical squamous cells of undetermined significance. Large cytology laboratories, providing over half of US cervical smear results, encourage reflexive human papillomavirus testing for such cases. Despite this, a study by Rebello et al in this issue urges "caution in the clinical use of testing for human papillomavirus testing" (p 893).[3] How do we interpret these somewhat conflicting messages about such an important issue in cervical cancer prevention?

Papers p 893

Methodological differences may account for some of the inconsistencies between study results. The recent US reports (the Kaiser Permanente Borderline Pap Study[1] and baseline data from the ASCUS-LSIL triage study or ALTS[2]) have involved women with a report of atypical squamous cells of undetermined significance on cervical screening. However, Rebello et al studied women with "persistent mild dyskaryosis or borderline nuclear change," which may represent *persistent* atypical squamous cells of undetermined significance or low grade squamous intraepithelial lesions (LSIL). Rebello et al's reported prevalence of human papillomavirus of 61% is not unexpected for a mixture of atypical squamous cells of undetermined significance and low grade squamous intraepithelial lesions, based on results from previous studies with the Hybrid Capture (HC II) test for high risk human papillomavirus types (Digene, Inc).[1 2 4]

However, Rebello et al's 35% prevalence of histologically defined high grade cervical intraepithelial neoplasia is much higher than the 7-16% reported previously for women with cytological findings of atypical squamous cells of undetermined significance or a low grade squamous intraepithelial lesion.[1 2 5] As suggested by Rebello et al, the requirement for *persistent* cytological abnormalities in their study population may account for higher disease prevalence. However, the disease prevalence may also be inflated by two additional differences in study methods.

Firstly, the Kaiser Permanente and the ALTS investigations obtained final diagnoses by colposcopically directed biopsy or endocervical curettage, while Rebello et al performed large loop excision of the transformation zone on all patients. Perhaps lesions that require a large loop specimen for detection contribute to the additional disease prevalence in the current study. Secondly, the Kaiser Permanente and the ALTS investigators defined high grade cervical intraepithelial neoplasia by consensus review by expert pathologists, while Rebello et al used a single pathologist to make the diagnosis. The issue at stake in each of these investigations is the diagnostic performance of the HC II human papillomavirus test in identifying women with histologically defined high grade cervical intraepithelial neoplasia. Thus, the accurate classification of this histological endpoint is paramount to defining test sensitivity, disease prevalence, and predictive value.

The Kaiser Permanente study of cases of atypical squamous cells of undetermined significance showed that human papillomavirus testing identified 89.2% of women (mean age 38 years) with underlying high grade disease; and in women aged under 30 the sensitivity and negative predictive value were 100%. Similarly in the ALTS trial the HC II human papillomavirus test in cases of atypical squamous cells of undetermined significance (mean age 29 years) was 95.9% sensitive in identifying women with underlying high grade cervical intraepithelial neoplasia, and the predictive value of a negative test was 98.9%. In the women studied by Rebello et al (mean age 32 years), HC II human papillomavirus testing had a sensitivity of 93% for high grade cervical intraepithelial neoplasia. The authors emphasise the modest negative predictive value (89%) in women aged under 30. This low negative predictive value, and the high disease prevalence, are inconsistent not only with the ALTS and Kaiser Permanente reports but also with previous human papillomavirus triage studies.[6 7]

We should be intrigued by high grade cervical intraepithelial lesions that are apparently not associated with high risk human papillomavirus. They may

BMJ 2001;322:878–9

be transient lesions associated with low risk human papillomavirus types, and thus not issues in terms of cancer risk. In addition, these lesions may represent likely candidates for reclassification on pathology review. Perhaps they are more accurately described as mild to moderate cervical intraepithelial neoplasia, as opposed to high grade lesions.

Alternatively, these may be bona fide high grade lesions with potential to progress. The cases were women with persistent *mild* abnormalities on cervical smear tests. This criterion may have selected women with lesions that were anatomically difficult to sample or identify, whether by cytology, human papillomavirus testing, or colposcopy. If that is the case, the study's implications may be more a treatise on the failings of colposcopy and biopsy than on human papillomavirus testing to identify occult lesions.

It would be informative to investigate further the specimens collected by Rebello et al, perhaps restricting the analysis to women who had atypical squamous cells of undetermined significance. Consensus review of the histology and subsequent recalculation of the HC II diagnostic performance would be useful. Comprehensive (including low risk types) human papillomavirus testing of confirmed high grade cases of cervical intraepithelial neoplasia that were negative for high risk human papillomavirus might solve another piece of the mystery. (Testing could be conducted on the lesion tissues if the original human papillomavirus specimens are not informative.) This venture would complement the forthcoming prospective data from the ALTS trial,[8] which will address whether dangerous precancerous lesions are lurking in women who were considered at low risk on the basis of a negative HC II human papillomavirus test.

M Michele Manos *senior investigator*

Division of Research, The Permanente Medical Group, 3505 Broadway, Oakland, CA 94611, USA

MMM has received research support and honorariums for speaking from Digene Inc.

1 Manos MM, Kinney WK, Hurley LB, Sherman ME, Shieh-Ngai J, Kurman RJ, et al. Identifying women with cervical neoplasia: using human papillomavirus DNA testing for equivocal Papanicolaou results. *JAMA* 1999;281:1605-10.
2 Solomon D, Schiffman M, Tarone B for The ALTS Group. Comparison of three management strategies for patients with atypical squamous cells of undetermined significance (ASCUS): baseline results from a randomized trial. *J Natl Cancer Inst* 2001;93:293-9.
3 Rebello G, Hallam N, Smart G, Farquharson D, McCafferty J. Human papillomavirus testing and the management of women with mildly abnormal cervical smears: an observational study. *BMJ* 2001;322:893-4.
4 The ALTS Group. Human papillomavirus testing for triage of women with cytologic evidence of low-grade squamous intraepithelial lesions: baseline data from a randomized trial. *J Natl Cancer Inst* 2000;92:397-402.
5 Kinney WK, Manos MM, Hurley LB, Ransley JE. Where's the high-grade cervical neoplasia?: the importance of minimally-abnormal Papanicolaou diagnoses. *Obstet Gynecol* 1998;91:973-6.
6 Cox JT, Lorincz AT, Schiffman MH, Sherman ME, Cullen A, Kurman RJ. Human papillomavirus testing by hybrid capture appears to be useful in triaging women with a cytologic diagnosis of atypical squamous cells of undetermined significance. *Am J Obstet Gynecol* 1995;172:946-54.
7 Ferris DG, Wright TC, Litaker MS, Richart RM, Lorincz AT, Sun XW, et al. Comparison of two tests for detecting carcinogenic HPV in women with Papanicolaou smear reports of ASCUS and LSIL. *J Fam Pract* 1998;46:136-41.
8 Schiffman M, Adrianza ME for The ALTS Group. The ASCUS-LSIL Triage Study (ALTS): design, methods, and characeristics of trial participants. *Acta Cytol* 2000;44:726-42.

Any casualties in the clash of randomised and observational evidence?

No—recent comparisons have studied selected questions, but we do need more data

Randomised controlled trials and observational studies are often seen as mutually exclusive, if not opposing, methods of clinical research. Two recent reports, however, identified clinical questions (19 in one report,[1] five in the other[2]) where both randomised trials and observational methods had been used to evaluate the same question, and performed a head to head comparison of them. In contrast to the belief that randomised controlled trials are more reliable estimators of how much a treatment works, both reports found that observational studies did not overestimate the size of the treatment effect compared with their randomised counterparts. The authors say that the merits of well designed observational studies may need to be re-evaluated: case-control and cohort studies may need to assume more respect in assessing medical therapies and largescale observational databases should be better exploited.[1 2] The first claim flies in the face of half a century of thinking, so are these authors right?

The combined results from the two reports indeed show a striking concordance between the estimates obtained with the two research designs. A correlation analysis we performed on their combined databases found that the correlation coefficient between the odds ratio of randomised trials and the odds ratio of observational designs is 0.84 ($P < 0.001$). This represents excellent concordance (figure). In fact, it is better than that observed when the results of small randomised trials and their meta-analyses were compared with the results of large randomised trials.[3] To complicate matters, the concordance has been worse when the results of specific large randomised trials on the same topic were compared among themselves.[3] Concato et al further observe that, for the five clinical questions they evaluated, observational studies for each question had very similar odds ratios between themselves,[2] whereas the results of the randomised trials were often very heterogeneous. Popular wisdom has it that a "gold standard" method should give more or less the same results when repeated several times, while a poor method would suffer from lots of variability. So should observational studies be the gold standard instead of randomised trials?

Such a thought would be anathema to most clinical trialists.[4] A closer inspection of the data suggests several caveats. Firstly, in six of 25 comparisons the

BMJ 2001;322:879–80