Discussion

Discussion of methods

Problems of developing an instrument using consensus statements

During the process of setting up the consensus group and obtaining permission from all the paediatricians in the region, one of the main objections to the study was the concern about the implications of the study and the manner in which the results would be used. The initial phase of the study was carried out soon after the implementation of the NHS reforms and there was considerable suspicion and fear about the increasing role of management in limiting clinical freedom. In particular, there was a concern that managers would use the development of criteria for the appropriateness of admissions to restrict the admission of children to hospital if they did not fulfill the criteria.

Part of the impetus towards the development of tools for the assessment of appropriateness of medical procedures was the need to contain costs. Insurance companies in the United States were keen to contain costs by defining circumstances in which hospitalisation was appropriate and restricting payment to only appropriate admissions [1,2]. The development of the internal market and competition amongst trusts introduced as a result of the health reforms in 1988, may still result in purchasers defining circumstances in which medical procedures and hospitalisation are appropriate. However, at present NHS purchasers and providers have not gone down this road, though the concern of paediatricians may not be misplaced.

The initial reluctance of clinicians to be involved in this exercise may have resulted in a more conservative development of criteria during the consensus process. This was apparent during the development of criteria for social admissions with clinicians keen to emphasis the uniqueness of the British child health care system in its treatment of social problems. Although criteria suggesting social admissions were included in the modified PAEP by consensus agreement, the issue of whether children with "social problems" need to be in hospital was never really discussed and challenged. In effect, the methods used to develop the PAEP merely provided a refined way of recording conventional wisdom about the efficacy of hospital intervention. Consensus statement therefore tend to err on the side of caution and in the view of some commentators[3] merely

re-inforce existing beliefs and prejudices. The influence of this conservatism on the results will be discussed in more depth later in this chapter.

A formal method for developing a consensus statement was not used for several reasons.
The use of more formal methods of consensus development such as the Delphi technique or modified Glaser technique[4] was considered and could have been used if there was difficulty in obtaining consensus. The advantages of modifying an instrument which had already been developed was that broad agreement had already been achieved by a group of clinicians, albeit working in a different country. There was therefore very little disagreement in the consensus method used on the majority of criteria which had been developed. The Delphi technique would have been more appropriate if a larger number of panelists were used and may still be necessary, if for example, the modified PAEP was to gain acceptance as a UK - wide instrument. However, for the purposes of the pilot and field study, the technique used to obtain consensus was successful in achieving its objective.

Limitations of validity testing

The greatest limitation to the use of the AEP was the concept of validity and determining the gold standard. The term validity refers to the extent that a measure agrees with truth, or at least another measure that is trusted to accurately measure the target variable [5,6]. In the case of the AEP there is no absolute indicator of whether the admission of episode of inpatient care is needed. In the absence of such a "gold standard" Gertman and Restuccia[1] used concepts such as "predictive validity" to measure the ability of a nurse reviewer to "predict" the appropriateness decision that would be made by an expert physician reviewer. They also felt that frequent use of the overrides would undermine validity and used the arbitrary cut off point of 10% as an indicator of good or poor validity. The problem of validity was addressed by claiming that face validity was assessed through critical review by physician committees in four different areas.

The literature is frequently confusing on the use of terms such as reliability and validity[6] and in instruments like the AEP where there is no gold standard the issue becomes more confused when researchers develop and use terms such as predictive validity in the wrong context. This is exactly

what Gertman and Restuccia[1] try and do and hence avoid one of the major criticisms of the PAEP - namely the absence of a gold standard.

My feeling is that in the absence of a truly valid measure, for example a gold standard, face validity is the only form of validity that can be commented on[5,6]. Attempts to define validity in terms of the ability of clinicians to be consistent in their use of the PAEP and the comparison of clinicians and nurses are really measures of reliability of two different groups and do not give the instrument any extra validity. Similarly, attempts to define predictive validity on the basis of nurses using the instrument prospectively and then assessing whether the admission was appropriate when the same instrument was used by clinicians again does not deal effectively with the absence of a gold standard .

Strumwasser and colleagues[7] attempted to use the majority opinion of clinicians as a gold standard. However, there are several problems with this approach. The early studies using the AEP showed that clinicians using subjective judgments to assess records had poor reliability when compared to clinicians using subjective criteria[1,8]. My study confirmed this finding with clinicians using subjective judgments achieving Kappa values of 0.34 compared to Kappa of 0.84 for clinicians using objective criteria. Consequently, using as a test of validity the views of clinicians whose reliability is poor, even under experimental conditions, is questionable. In addition the failure of the AEP to ascertain cases that should have been admitted but were not (false positives) limits the usefulness of such measures[5].

The fact that a group of clinicians and trained reviewers can reach high levels of reliability is a strength of the instrument and could give it added credibility amongst clinicians, but in terms of its use as a research instrument or audit tool, the reliability achieved by the trained reviewers is probably the most important factor.

The high Kappa values that were obtained by the trained reviewers in the pilot study and their consistency in their reliability during the filed study was almost certainly the result of the extra training the reviewers received with very specific instructions on the use of the instruments. During the pilot study, several physicians had problems using the PAEP because there was a greater tendency to impose subjective judgments over the criteria rather than use the override function. The reluctance of several clinicians not to use the manual in difficult judgments (as an oversight rather than a conscious decision) created some minor problems. For

example, although criteria for admission were present, some clinicians failed to register this. During the pilot study, this was overcome by the author reclassifying the judgment and using it as an additional opportunity to train the reviewers in the correct use of the instrument. My experience with training of the reviewers confirmed my own impression that training was a key element in the effective use of the instrument.

The short comings of the instruments such as the AEP should therefore be recognised and considered as part of their limitations. In the case of the PAEP all that can be said is that it was developed by the consensus of a group of physicians which gave it face and content validity i.e. it was felt by the panel that the instrument could be used to measure appropriateness of care and that the criteria were sufficient to do this. This instrument when applied in field settings had high reliability and could be used by trained researchers with consistent results.

The PAEP as a clinical guideline

There are a number of other intrinsic problems associated with the use of instruments such as the AEP. They only study interventions that have already occurred (ignoring issues of inappropriate failure to perform an intervention), they ignore patient preferences[9] and as explained earlier, they only reflect the consensus of experts, who may have little clinical science on which to base their judgments. Looking at Anderson's conceptual model of admissions[10], it only quantifies one of the key elements (Medical Care Factors) associated with admissions. However, as an instrument for measuring medical care factors associated with admissions, it remians useful, valid and reliable.

At one level, the AEP can be considered a clinical guideline defining a set of criteria for admission to hospital. Grimshaw and Russell[11] have defined desirable attributes of clinical guidelines and it is worth assessing the AEP against these criteria.

These are defined as follows:

Table 1

Attributes considered essential for the development of clinical guidelines [12]

Attribute	**Meaning**
Validity	Correctly interpreting available evidence so that when followed valid guidelines lead to improvements in health
Cost effectiveness	Guidelines lead to improvements in health at acceptable costs
Reproducibility	Given the same evidence, another guideline group produces similar recommendations
Reliability	Given the same clinical circumstances another health professional applies them similarly
Representative development	All key disciplines and interests (including patients) contribute to guideline development)
Clinical applicability	Target population is defined in accordance with scientific evidence
Clinical flexibility	Guidelines identify exceptions and indicate how patient preferences are to be incorporated in decision-making
Clarity	Guidelines use precise definitions, unambiguous language, and user-friendly formats.
Meticulous documentation	Guidelines record participants, assumptions and methods and link recommendations to the available evidence
Scheduled review	Guidelines state when and how they are to be reviewed
Utilsation review	Guidelines indicate ways in which adherence to recommendations can be sensibly monitored.

Each of these attributes will be now considered individually to ascertain whether the PAEP fulfills some of the minimum criteria for it to be considered as a clinical guideline.

Validity

This has been discussed in the previous section. The development of the PAEP was not based on available

evidence but on the consensus views of clinicians. There is for example no evidence that many of the criteria are associated with a negative outcome unless admitted to hospital. For some criteria, for example coma requiring hospitalisation, there is probably sufficient agreement that hospitalisation is the most appropriate intervention but in others for example admission for social reasons there is almost no evidence suggesting that hospitalisation produces the best outcome.

Cost effectiveness

There is no evidence on the cost effectiveness of the PAEP. It has been used by some insurers to try and define when a payment is made for hospitalisation but the instrument was not developed for this purpose.

Reproducibility

This has yet to be tested with the PAEP and would require another group of paediatricians to produce similar criteria in the UK. This is an area for further work if the PAEP is to be used widely as an audit instrument.

Reliability

This has been assessed extensively.

Representative development

Patient views have been excluded in the development of the AEP. Hospitalisation is dependent on the view of the referring GP, the admitting doctor and the patients (parents in the case of children.) At present there is no mechanism in the PAEP to define parents views and this may be an important defect in such an instrument. The absence of patient views confirms my reservations about the use of the PAEP as an instrument to define appropriate admissions rather than as an audit tool and utilisation review instrument. The views of nurses have not been taken into account and potentially could be another serious omission if the PAEP is to be used more widely. However a health visitor was employed as a reviewer on the project and several modifications were made to the criteria based on comments and suggestions from this reviewer.

Clinical applicability

The PAEP fulfills this criterion

Clinical flexibility

The use of overrides allows a degree of flexibility on the part of the reviewer and therefore fulfills this criterion though patent preferences are not included as described above.

Clarity

The PAEP fulfills all these criteria

Meticulous documentation

The PAEP fulfills this criterion because the process of development was explicit. The use of the trainer manual also defined the methods of application.

Scheduled review and Utilisation review

Not applicable

It can therefore be seen that using well accepted criteria for assessing clinical guidelines, the PAEP fulfills very few criteria. This further emphasises the point that the PAEP cannot be used as a prospective guideline for assessing admissions and its strength remains in its use as a utilisation review instrument.

Other methodological problems with the PAEP

In a recent article Phelps[13]made some serious criticisms about the use of measures of appropriateness and whilst most of these relate specially to methods related to the use of medical interventions and the Rand approach , some of the criticisms are also applicable to the PAEP and need to be considered in more detail.

Firstly, Phleps argued that instruments such as the PAEP only study interventions that have already occurred (hence ignoring issues of inappropriate failure to perform an intervention), they ignore patient preferences and they only reflect the consensus of experts, even though they are based on clinical/physiological statements. This is best illustrated with the condition of asthma.

The AEP lists several clinical and nursing interventions which might be relevant to asthma. This includes a rapid pulse, respiratory distress and frequent use of a nebuliser. There are occasions when a patient is positive for all these criteria and is still not admitted, parents may express a need for admission because of their own anxieties and sometimes express a desire for admission. The later are not accounted for in any of the criteria and ideally if any judgments are to be made about inappropriate admissions, then a knowledge about admissions that should have been hospitalised but were not is also important (false negative admissions).

Phelps also considers other methodological weaknesses which can be summarised as follows.

If the AEP can be considered as diagnostic test, then like any diagnostic test, there can be two types of error. .

True appropriate admission	**False positive** (appropriate but should have been inappropriate)
False negative (should have been an appropriate admission but classified as inappropriate)	**True inappropriate admission**

Because of the absence of a true gold standard, there may be biased estimates of the rate of inappropriate admissions. The process can label a treatment as inappropriate either correctly, when it was inappropriate, or incorrectly, when the doctors proceeded appropriately but the method mislabeled it as inappropriate. False positive results which classify treatments as appropriate when they were inappropriate occur at the rate of 1- specificity and false negative results (classifying a treatment as inappropriate when they were appropriate) occur at the rate of 1-sensitivity. The estimated rate is therefore:

estimated rate = true rate x (sensitivity) + (1-true rate) x (1-specificity).

estimated rate/true rate = sensitivity + (1-true rate)/true rate X (1-specificity)

The AEP like any diagnostic test will have a sensitivity and specificity. Because of the lack of a gold standard, I have not been able to determine this. However, the actual inappropriate rate will be dependent on the true prevalence of the inappropriate rate, and if it is low, the false positive will outnumber the true positives. Phelps goes on to argue that the estimated rate understates the true rate only when the specificity is quite high, and the degree of understatement, when it occurs is relatively small. The estimated rate overstates the true rate as the specificity falls: this implies that the better the practice (i.e. the better the true appropriate admission rate, then the larger the error in the estimates of inappropriate admission).

Phelps's criticism is directly relevant to my findings because the level of inappropriate admissions is quite low. Using his calculations and assuming the PAEP has a sensitivity and specificity of 95 per cent then it is possible that the estimated rate is double the true rate in some hospitals where the inappropriate rate is lower than 4 per cent. This can be

explained in more detail using the above equations to calculate the estimated rate of admissions.

Assume 1000 patients assessed and 4 per cent inappropriate rate. The PAEP would correctly identify 38 out of 40 cases which were truly inappropriate and falsely identify 48 out of 960 appropriately admitted cases as inappropriate. The estimated overall inappropriate rate would be 0.086 and the ratio of the estimated true rate would be 2.15 (0.086/0.04).

As with all diagnostic tests, the predictive value of the positive result depends heavily on the underlying rate of occurrence of the event measured by the test. False positives commonly outnumber true positives when the underlying rate of occurrence is low. Only if the PAEP was a perfect diagnostic test would this problem disappear.

Phelps goes on to demonstrate the relationship between the estimated and true rate using different sensitivities and specificity's and concludes that the estimated rate understates the true rate only when the specificity is quite high and the degree of understatement when it occurs is relatively small . The estimated rate overstates the true rate as the specificity falls but the problem gets worse as the true rate falls. In the cases of my findings there is the suggestion that better actual practice leads to larger relative errors in the estimates of the inappropriate treatment.

Phelps does of course point out that there are no available estimates of the sensitivities and specificity's of instruments such as the PAEP but suggests ways that such figures might be obtained using emerging statistical techniques. However the present state of the art is such that there significant problems with the use of instruments such as the PAEP particularly if people use them as gold standards. However, it does not negate their use as research instruments or audit tools - a theme to which I will be returning later in this chapter.

Reliability of the modified PAEP

The agreement between the consensus panel using the modified PAEP admission criteria was very high (K= 0.848) and highly significant. Between the trained researchers, agreement was even higher (K=0.948). Agreement between the consensus panel on the Day of Care criteria was moderate though remained high amongst the researchers. I can not readily explain the reason why there was greater inter-rater agreement for the Admission criteria compared to the Day of Care criteria though interestingly, amongst the trained researchers, agreement for both the Admission criteria and Day of Care criteria was excellent. This therefore may reflect the difference in the level of training because for the clinicians there seemed to be greater interest in the Admission criteria than in the Day of Care criteria. The inappropriate rate for the Day of Care was much higher in the pilot study (47%) as opposed to 37% for the admission rate and the lower inter-rater agreement may be related to this. However, the grading frequently quoted for the Kappa statistics suggests that, as far as the Day of Care criteria are concerned, agreement is moderate and still significantly greater than chance.

The maintenance of consistent reliability between the researchers during the field study was reassuring and suggests that the results across all hospitals are comparable and not subject to large difference between the raters.

In summary, it can be argued that there are serious methodological flaws with the use of the PAEP particularly if an attempt is made to use it as a gold standard by which to assess admissions. This directly compromises attempts to define a truly valid measure against which to assess it. However, its reliability is very high, both when used by clinicians and trained reviewers. The stability of this measure during prolonged field testing suggests that, with appropriate training of reviewers, it remains a highly reliable utilisation review instrument.

Discussion of results

Sample

There was a large variation in the sampling percentage between the different districts and this probably reflects differences in estimating the denominators. Denominators were obtained from the summary HES data provided by the region and close examination of this data showed large variation in consistency of recording. For example in some districts normal births were included in the paediatric admission totals and it was not always possible to exclude these prior to records being selected by the researchers. This was certainly the case at both Queen Mary's Carsholton and Kingston hospital where the percentage sampled appears to be the lowest. Overall however, the researchers achieved their objective of attempting to sample approximately 10% of admissions excluding routine surgical admissions, admissions to burns units and rehabilitation wards.

The inconsistency of HES data in recording normal healthy babies and known problems in trying to assess the available beds in paediatrics[14] makes it difficult to use HES data to determine denominators and look at the relationship of appropriateness to the total number of admissions and bed availability. Any attempt to investigate the variation in appropriate admission rates will need a detailed investigation of the link between appropriateness and bed availability. This will only become possible when there is a consistent means of measuring bed availability and occupancy at a national level.

Age and gender

As stated in the results section, the difference which I found between appropriate admissions for males and females does not have a reasonable explanation. The finding could be a chance finding. Clinically it is of no significance. There is a significant age distribution across hospitals which can not be readily explained. The random sample selection should have excluded any systematic bias. In view of the variation in age between hospitals, this is controlled for in analyses where it may be important.

Length of stay

As expected there was a large variation in length of stay between the different hospitals. It is almost certainly an important variable associated with appropriateness. The variation in LOS between hospitals

could also be accounted for by different policies which were operating administratively. For example, hospitals which had no admissions less than 24 hrs probably had a different way of accounting for children admitted and discharged on the same day. Alternatively some hospital consultants may have unwritten policies stating that no admission should be discharged unless reviewed by a senior member of staff. This would result in fewer admissions being discharged within 24 hrs.

This variation in administrative policies has not been formally assessed in this study. There is for example much greater consistency between hospitals in admissions where lengths of stay were between 25-48 hrs. Variations in lengths of stay greater than 4 days may be unduly biased by individuals who had exceptionally long length of stay. Traditionally, this is taken into account by quoting median lengths of stay but for the purposes of this study I was only interested in the appropriateness of the day before discharge and the admission. Variations in lengths of stay between hospitals although interesting in their own right are not the purpose of this study, though it should be pointed out that much of the literature in variation of medical practice initially looked at the variation in lengths of stays between hospitals. Interestingly, 58% of the sampled admissions took place during 'on-call' hours, suggesting that unwritten policies may be operating in some units concerning discharge of these patients.

Appropriateness of admissions and Days of Care

A review of the literature prior to the beginning of the study had led me to expect a much higher level of inappropriate admissions than the level of 8% that I found in South West Thames. There had been no previous studies of admissions in the UK using the AEP and work carried out by other researchers claiming to use instruments similar to the AEP[15] were criticised because of deficiencies in methodology. In their study of admissions using the Oxford Bed Study Instrument, Victor and Khakoo[15] reported an inappropriate rate of admission of 1%. However, as pointed out by several researchers, the Oxford Bed Study Instrument could not be directly compared to the AEP because it did not treat the day of admission separately from the subsequent days in hospital. This would inevitably lead to bias and an estimation of a lower inappropriate admission rate. The use of the Paediatric AEP in this study closely followed the method described by Restuccia, Gertman and Kemper[1,8] who reported an inappropriate admission rate of 15% in the group of hospitals they studied. The large differences in health care systems between the United States and the United Kingdom means that rates of inappropriate

hospital use are not directly comparable but do raise questions about the possible reasons for the differences between the two countries.

One of the reasons for the significant differences between different hospitals could be the possibility of systematic bias as a result of the reviewers use of the instrument. There appeared to be no relationship between the rates of inappropriate admission and an individual reviewer. An assessment of the null hypothesis that there was a difference between the reviewers classification of an admissions using McNeemars test was not proved.

Having excluded the possibility that systematic bias was not a reason for the differences between the hospitals, I need to consider several other reasons. As discussed earlier, there is no gold standard of appropriate admissions and it is possible that in some of the units with very low inappropriate admission rates, there are some patients who may need to be admitted but are not. An alternative explanation is that these units have the best clinical practice and other units should try and achieve results similar to them. It is interesting to note that two of the three hospitals with high rates of inappropriate admissions (St George's and Queen Mary's Carsholton) are teaching hospitals and partly act as tertiary referral centres within the region. It is not clear why this should result in a higher level of inappropriate admissions because I would expect that staffing ratios between senior and junior staff would be amongst the best in the region. Furthermore, junior staff would more likely be following careers in paediatrics and hence be more qualified than staff in district general hospitals. There also appears to be no relationship in inappropriate admission rates between rural and inner city hospitals - for example East Surrey and Worthing (rural hospitals) have higher inappropriate admission rates than Epsom and Frimley (rural hospitals) and Mayday and Queen Mary's Roehampton (inner city hospitals).

I have no data on the quality or extent of primary care services in the areas that I studied and one area of further study that would be useful would be to investigate the relationship between appropriate admissions rates and the level and quality of primary care services. This latter relationship may be critical to an understanding of why there is a lower inappropriate admission rate in the UK than in the United States. There is some evidence that higher levels of primary care services are associated with lower hospitalisation of children[17], and my findings that admissions via the general practitioner tend to be more appropriate than admissions via

casualty may be a pointer in this direction. This will be discussed in more detail later.

Although the variation between hospitals in the admission criteria was highly significant, the overall low rate of inappropriate admissions raises the issue of whether this variation is clinically important or not. In any study of 13 hospitals and 3000 admissions, almost any variation is likely to be statistically significant.

However, the variation in day of care criteria was much greater (inappropriate rate varied from 18% - 65%) and is clinically more important. As mentioned previously, day of care criteria was assessing the day before the discharge of the patient in any admission greater than 48 hrs. My findings suggests that a large proportion of these days in hospital are inappropriate. Although there has been an trend towards decreasing lengths of stay for paediatric admissions, my findings suggest that there is still considerable room for improvement from earlier discharge. The fact that some units are achieving inappropriate rates of days of care of 18-30% (Frimley, St peters and RSCH) compared to rates of 60-65% in other units (Queen Mary's Roehampton, Crawley, Worthing) suggests that there is considerable scope for improvement. The very nature of the day of care criteria suggest that much of this continuing care in hospital could be provided in a different setting - for example the child's home. Improving administrative procedures as outlined by the Audit Commission [18] may have significant impact in reducing lengths of unnecessary days in hospital. The use of an instrument such as the PAEP would provide a useful measure for clinicians in auditing and monitoring progress of their units towards levels of inappropriate care achieved by the better units.

Use of overrides

Gertman and Restuccia[1] placed a large emphasis on the use of overrides as a means of monitoring the reliability of the AEP and keeping a check on its validity. Table 17 confirms that the use of overrides was well below the arbitrary limit of 10% set by Gertman and Restuccia[1] and that it remained low throughout the length of the study. In itself, the low percentage of overrides does not give the PAEP any extra validity but confirms that there were very few instances in the medical records when the criteria were not sufficient to make a judgment.

Factors associated with inappropriate admissions

Rather than trawl through the data to ascertain factors associated with inappropriate admissions, I determined some a priori hypotheses based on existing knowledge about some of the factors that may be associated with inappropriate admissions. In addition to the a priori hypotheses that I tested, I am also aware that there is considerable scope for testing additional hypotheses particularly looking at relationships between inappropriate admissions and social factors and published figures for hospitalisation rates. My reluctance to test these hypotheses was based on the quality of the available data. For example, although I collected information on the addresses of all admitted children with a view to linking this with census data and hence make links with social factors associated with inappropriate admissions, I felt that problems with the ecological fallacy would make it very difficult to draw meaningful conclusions.

Similarly, looking at the relationship between inappropriate admissions and hospitalisation rates published annually by the Department of Health might be interesting, However, I was aware as a result of analyses of HES data provided to me by the RHA (which are then reported to the DoH) that there are significant problems associated with this data. Although the DoH is trying to make sure that only paediatric admissions are counted several districts in South Thames were still having well babies counted in their totals for paediatric admissions. There are also problems associated with the use of denominators and cross boundary flows, particularly in inner London districts. Therefore although the relationship between inappropriate admissions and hospitalisation rates is worth examining the present scope of the study and the limitations of the existing data sets convinced me to leave this analysis to a later stage.

The a priori hypotheses that I developed were therefore looking at the association with age, length of stay, time of admission and source of referral.

The relationship between inappropriate admissions and age

Table 11 shows that overall, the younger the age of the child the more likely it is that the admissions are inappropriate. This is not a surprising finding and is almost certainly related to the uncertainty associated with the admission of young children under 1 to hospital. Physical signs in young children are vague, the history is frequently not very specific and the ability to exclude serious illness is difficult. It is therefore not surprising that more children under one are admitted to hospital and that more of these admissions are classified as inappropriate.

If younger children contribute disproportionately to the overall inappropriate admission rate are they any strategies that may reduce this rate. Although clinical assessment in this group is difficult, serious illness and mortality is relatively rare. Parental and general practitioner reassurance may therefore be a large factor in determining whether a child is referred to hospital. Assessment by a junior doctor, frequently less experienced than the referring general practitioner compounds the problem and the child is admitted to hospital.

There are probably several strategies to deal with this complex interaction of factors and include parental education about dealing with minor illnesses in childhood. Better support for general practitioners so that an ill child can be assessed more frequently and admission policies that require assessment of an admission by a more senior doctor are interventions that may improve the appropriateness of admissions. The PAEP offers a means of assessing and evaluating such interventions.

The relationship between inappropriate admission and length of stay

Table x shows that the shorter the length of stay the more likely the admission was assessed as inappropriate.

The large proportion of admissions less than 24 hrs which were judged to be inappropriate probably reflects the clinical uncertainty referred to above. It is likely that hospitals which showed no inappropriate admissions in this category either had a policy to ensure that no child was admitted for less than 24 hrs or more likely, the hospital administration system was unable to record admissions for less than 24 hrs because no administrative mechanism existed for this. Many hospitals have day assessment units whereby children can be admitted for a few hours for investigation to exclude serious illness or allay parental anxiety and these

are frequently counted as hospital admissions lasting less than one day, though in some units they do not count as a hospital admission.

However those hospitals that did admit children for less than 24 hrs showed that a large proportion of these admissions were classified as inappropriate (range 16% - 48%). This variation was less marked in lengths of stay of 1-2 days and virtually disappeared in lengths of stay greater than 3 days. Almost no admissions were classified as inappropriate once the length of stay was greater than 3 days. There are several possible explanations for this finding.

The natural history of most minor childhood illnesses is that it is of short duration. However, clinical uncertainty concerning the diagnosis of these illnesses, particularly in the earlier stages of the illness results in increasing admissions. Because the illness resolves very rapidly, it soon becomes apparent that the illness is not serious and the child is discharged home. The initial decision to admit is therefore not based so much on the severity of the child's illness but on the uncertainty of the referring clinician. Using an instrument such as the PAEP which bases its criteria on the patients clinical condition and the intensity of the service required will result in minor illness not fulfilling the criteria as often as more serious illness. Hence more of the admissions with short lengths of stay are classified as inappropriate. The converse of this reasoning is that children who stay in hospital more than three days are usually unwell and require longer lengths of stay.

The challenge therefore in reducing inappropriate admissions is to introduce changes which help reduce clinical uncertainty in minor illness and hence reduce the need for hospitalisation. However, it could be argued that the present organisation of services may in fact be doing this very well because based on the PAEP criteria most children with minor illness are discharged rapidly and inappropriate admissions do not result in longer lengths of stay. However, the cost effectiveness of this approach needs to be assessed in more detail.

Relationship between appropriateness of admissions and time

Overall there was no significant difference between the assessment of admission and the time of day or day of week. I had expected more weekend admissions and admissions during on call hours to be classified as inappropriate on the basis that more junior staff were responsible for the decision to admit during these hours. Interestingly, in some hospitals,

more admissions during weekdays were inappropriate compared to weekends. The absence of a difference is probably due to the fact that it is generally always the same grade of junior staff that makes the initial decision to admit with the decision being reviewed at a later time by a more senior colleague. The difference between hospitals for inappropriate admissions between weekdays and weekends was not significant when assessed by the test for the homogeneity of the odds ratio. The differences in inappropriate admissions between hospitals between daytime and on call just reached statistical significance.

Relationship between source of referral and appropriateness of admissions

I had hypothesised that admissions made via the general practitioner were more likely to be inappropriate compared to admissions via the Accident & Emergency department. Overall my findings suggested that there was no difference in the assessment of admission and source of referral. This could be explained by several factors. Depending on the policy of individual hospitals, admissions to hospitals may have been directed via the A & E department and classified as such even though they originated from a general practitioner. There is no means of ascertaining this from the hospital records that we reviewed. Many GP admissions that were classified as inappropriate may have been directed through this route though it is unlikely that this would have resulted in systematic bias. What is certain is that admissions via the A&E department would have been assessed by a paediatrician (albeit a junior doctor) prior to admission to the ward and that more diagnostic facilities would be available to reduce clinical uncertainty in decision making. It is therefore not unreasonable to expect that inappropriate admissions via this route would be significantly different compared to referrals directly from GP's. The fact that my results failed to demonstrate this suggests that the level and quality of primary care services may alter the hospitalisation rate. The evidence for this in the United Kingdom is mixed[19,20]though it has been postulated as a reason for differences in variations of hospitalisations in parts of the United States[17].

GP's are uniquely placed to know more about the home circumstances of children than hospital physicians and this knowledge may influence the decision to hospitalise a patient. Because GP's are more likely to see minor self limiting illnesses, they are better able to judge the severity of these illnesses and more importantly place them in a social context. They would also probably be better judges of serious illness partly because they

would have more experience than the most junior member of the hospital paediatric team. The fact that more GP referrals stayed in for longer lengths of stay, suggesting that these children had more serious illnesses supports this observation.

As discussed earlier, inappropriate admissions were concentrated more frequently in children under 1 and reasons for this observation have already been discussed. Interestingly, fewer referrals via GPs in this age group were classified as inappropriate compared to referrals via the A & E department. This again suggests that clinical uncertainty amongst GP's for this group of children is much less than for referrals via the A & E department. Once again, greater knowledge of the child and its family and social background may reduce the need for admission and result in more appropriate admissions where this is deemed necessary.

Prevention of hospital admission with the use of pediatric home nursing teams and more intensive support by General Practitioners has been cited by the Audit Commission[18] as one way of reducing unnecessary hospitalisation and inappropriate admissions. The Nottingham Paediatric Community Nursing team [21,22] is cited as an example of an intervention which has reduced hospitalisation of children. However, it should be borne in mind that increasing care at home may increase the caregiver's burden and therefore the secondary effects of such interventions and their unintended consequences need to be carefully considered. My findings suggests that GP's are no worse than hospital A & E departments in determining the appropriateness of admissions and in specific instances are better in making judgments about the appropriateness of admissions. Better nursing support may further improve GP's ability to monitor and look after children at home and avoid unnecessary hospitalisation.

The fact that 21% of appropriate admissions (Table 26) were solely for the purposes of receiving services (mainly nursing) in hospital suggests that many of these services could be provided in the home - for example frequent monitoring by nurses and nebuliser use. 41% of cases staying in hospital more than 48 hrs were in hospital to receive nursing support services, many of which could be provided in the home setting. Proper trials will need to be carried out to assess the effectiveness and cost-effectiveness of such interventions. The use of the PAEP as a monitoring tool may be useful to monitor outcomes of interventions.

Discharge diagnosis and appropriate admissions.

There are considerable problems with discharge diagnosis and HES data from which information on discharge diagnosis for this study was obtained. Nearly 20% of diagnoses were missing and a further 23% were classified under the heading of symptoms, signs and other unspecified conditions. The possibility of systematic bias cannot be excluded when nearly 40% of admissions cannot be classified accurately. However, nearly one third of classified admissions were for gastro-intestinal infections, acute upper and lower respiratory infections and asthma. Within these four categories, 10% of admissions for gastro-intestinal infections were judged inappropriate as were 13% of admissions for upper respiratory infections and 9% for lower respiratory infections. In contrast only 3% of admissions for asthma were classified as inappropriate. Gastro-intestinal infections and upper respiratory infections are self-limiting minor illnesses in developed countries and mortality from these conditions is extremely low. Better management of these conditions in a primary care setting may reduce unnecessary admissions and once again , better support facilities in primary care could make a large reduction in hospitalisation for these conditions.

Summary and potential for further work

The PAEP developed as part of this study has been effectively tested in a field setting and important results have been obtained. Using the same methodology, it is likely that the study can be repeated in other areas of the United Kingdom.

Although there has been a suggestion both in published research and in government sponsored reports that inappropriate admissions are a problem in paediatrics, my findings do not support this view. However, there may still be scope to reduce further the length of stay of paediatric admissions with improved discharge planning as shown by my findings of high inappropriate Days of Care. What can be stated fiarly confidently based on the result of my study is that the appropriateness of providing some form of care is not in question. The issue is whether there should be other forms of care which may be cheaper, more efficient , less restrictive and more favourable from the patient's perspective than hospital admission.

The possibility of using the PAEP as an audit tool needs to be explored further. Feeding back results obtained form the use of the PAEP to clinicians and monitoring the effect on behaviour in terms of discharge planning and better admissions policies is an area for further research.

The use of the PAEP as a research instrument to monitor changes in practice following service interventions such as the implementation of home nursing teams needs to be developed. This is an area in the UK which needs to be explored in more detail and is likely to yield innovative developments particularly because of the already well developed primary care systems which operate in this country. Methods to improve support to general practitioners are an additional area for investigation with the PAEP being used to monitor the effects these interventions.

Bibliography (Discussion)

Ref ID : 204
1 Gertman PM, Restuccia JD. The appropriateness evaluation protocol: A technique for assesing unnecessary days of hospital care. *Med Care* 1981;**19**:855-871.

Ref ID : 297
2 Berwick DM. Health services research and quality of care. Assignments for the 1990s. *Med Care* 1989;**27**:763-771.

Ref ID : 1112
3 Pharoah POD, Alberman ED. Mortality of low birthweight infants in England and Wales 1953-79. *Arch Dis Child* 1981;**56**:86-89.

Ref ID : 260
4 Glaser EM. Using behavioral science strategies for defining the state-of-the-art. *J Applied Behav Sci* 1980;**16**:79-92.

Ref ID : 1122
5 Streiner DL, Norman GR. *Health measurement scales. A practical guide to their development and use.* Oxford: Oxford University Press, 1989;

Ref ID : 249
6 Carmines EG, Zeller RA. *Reliability and Validity assessment*. London: Sage Publications, Inc, 1979;18-19

Ref ID : 272
7 Strumwasser I, Paranjpe NV, Ronis DL, Share D, Sell LJ. Reliability and validity of utilization review criteria. Appropriateness evaluation protocol, standardized medreview instrument, and intensity-severity-discharge criteria. *Med Care* 1990;**28**:95-109.

Ref ID : 231
8 Kemper KJ, Fink HD, McCarthy PL. The reliability and validity of the pediatirc appropriateness evaluation protocol. *QRB* 1989;**March**:77-80.

Ref ID : 1123
9 Hopkin A. What do we mean by appropriate health care. *Qual Health Care* 1993;**2**:117-123.

Ref ID : 236
10 Anderson HR. The epidemiological value of hospital diagnostic data. In: Bennet AE, ed. *Recent advances in Community Medicine No.1*. London: Churchill Livingstone, 1978;175-193.

Ref ID : 1124
11 Grimshaw J, Russell I. Achieving health gain through clinical guidelines. I : developing scientifically valid guidelines. *Qual Health Care* 1993;**2**:243-248.

Ref ID : 1145
12 Anonymous Effective health care: Implementing clinical practice guidelines. 1994;**8**:

Ref ID : 287
13 Phelps CE. The methodologic foundations of studies of the appropriateness of medical care [see comments]. *N Engl J Med* 1993;**329**:1241-1245.

Ref ID : 1126
14 MacFaul R, Long R. Major problems with paediatric bed usage statistics? [see comments]. *Arch Dis Child* 1991;**66**:504-507.

Ref ID : 289
15 Victor CR, Khakoo AA. Is hospital the right place? A survey of 'inappropriate' admissions to an inner London NHS trust. *J Public Health Med* 1994;**16**:286-290.

00000000000000001146 ***** Not found in the database "ANEEZ"**

Ref ID : 211
17 Perrin JM, Homer CJ, Berwick DM, Woolf AD, Freeman JL, Wennberg JE. Variations in rates of hospitalisation of children in three urban communities. *N Engl J Med* 1989;**320**:1183-1187.

Ref ID : 1116
18 The Audit Commission for England and Wales. Children first: a study of hospital services. 1993;

Ref ID : 1132
19 Durojaiye LI, Hutchison T, Madeley RJ. Improved primary care does not prevent the admission of children to hospital. *Public Health* 1989;**103**:181-188.

Ref ID : 271
20 Rajaratnam G. A study of admissions to paediatric beds. *Postgrad med j* 1991;**67**:50-54.

Ref ID : 1136
21 Dryden S. Care in the community: the work of Nottingham Paediatric Community Nursing Team. *Paediatric Nursing* 1989;19-20. Ref ID : 1137

22 Dryden S. Care in the community: The work of Nottingham paediatric community nursing team. *Paediatric Nursing* 1989;17-18.