

A CRITICAL APPRAISAL OF THREE RESEARCH STUDIES RELATED TO PERIPHERAL VENOUS CANNULAE AND THE INCIDENCE OF PHLEBITIS.

INTRODUCTION

It is widely acknowledged that intravenous therapy is an integral part of total patient care of many hospital patients. (Goodinson 1990, Stonehouse 1996, Parker 1999). A variety of devices are used to deliver intravenous fluids, medication, total patient nutrition and blood products. The most common delivery method being by the percutaneous insertion of a peripheral venous cannula. According to Campbell (1998) twenty five million NHS patients per year are estimated to receive some form of intravenous therapy via a peripheral route.

Peripheral venous cannulation has been associated with sepsis that causes substantial morbidity in patients, prolonged hospitalisation and increased hospital costs. (Maki 1991 cited in Campbell 1998 p1308, Pearson 1996 cited in Scales 1997 pS4, Curran et al 1998). Three main theories appear to provide a basis for predicting the occurrence of phlebitis. Physical theories support the premise that veins are traumatised by physical contact with organic or inorganic materials, The principle of the septic theory suggests that phlebitis develops as a direct result of sepsis or infection. Whilst Chemical theories propose that phlebitis is caused by the irritation of the vein wall by infusion fluids. (Campbell 1998). However, the literature appears to suggest that most cannula-related infections result from the migration of skin organisms at the insertion site, eventually colonizing the cannula tip and entering the blood stream. (Kelsey and Gosling 1984,

Fletcher and Bodenham 1999). The complications of the use of peripheral cannulation can therefore, as Horton and Parker (1997 cited in Parker 1999 p1492) indicate, range along a spectrum of symptoms including, most commonly, phlebitis but also extravasation and other localised and systemic infections.

According to Pearson (1996 cited in Scales 1997 pS4), the length of time the cannula is in situ is an important factor in the development of sepsis, with the risk increasing the longer the cannula is in place. Pearson's (1996) research findings were used in the formation of the existing guidelines on peripheral venous cannulation within my Trust. These guidelines are now undergoing a review to ensure that they are in line with the Trust's Risk Management Strategy. As part of the Trust review process I am carrying out a literature review on the care of patients with peripheral venous cannulae in order to identify associated risk.

From a professional perspective, nursing staff are increasingly performing cannulation. Professional and personal accountability underpins nursing practice and Clause 6, section 6.5, of the Code of Professional Conduct (2002) refers to the nurses' responsibility in delivering care based on current evidence. The initial literature search that was carried out raised issues around the consensus of opinion regarding the length of time cannulae can be safely left in situ.

LITERATURE SEARCH

As Benton and Cormack (1996 cited in Cormack p 69 1996) acknowledge, conducting a literature search can be very time consuming. However, as Carnwell and Daly(2001) discuss, one of the key elements of a literature review, is a well-planned search and selection strategy. Several frameworks

are available for searching the literature and they can be useful tools for maintaining a focussed, systematic and thorough approach.

The literature search for this assignment was conducted by using CINAHL and MEDLINE databases. To establish the breadth of literature available, the key words used were: cannula, cannulation, catheter, infection, peripheral and time. Manual searches of the Journal of Advanced Nursing (1992 – 2002) and the British Journal of Nursing (1992 – 2002) were also carried out.

From the wider search, inclusion and exclusion criteria were used to narrow the focus onto the specific topic.

Inclusion Criteria

- ◆ Nursing perspective
- ◆ Acute/emergency care
- ◆ English language
- ◆ Research papers

Exclusion criteria

- ◆ Year of publication, limit to 1992-2002
- ◆ Medical models

With the criteria imposed the search resulted in retrieving fifteen articles.

This assignment will focus on a critical review of three primary research articles from the search of the literature relating to the relationship between peripheral venous cannula dwell time and the development of phlebitis.

ARTICLES FOR REVIEW

- ◆ CURRAN E.T., COIA J.E., GILMOUR H., McNAMEE S., HOOD J. (2000) Multi-centre research surveillance project to reduce infections/phlebitis associated with peripheral vascular catheters. **Journal of Hospital Infection. 46:194-202**
- ◆ HOMER L.D, (1998) Risks Associated with 72 and 96 hour peripheral intravenous catheter dwell times. **Journal of Intravenous Nursing 21(5); 301-305**
- ◆ LAI KWAN KEW (1998) Safety of prolonging peripheral cannula and IV tubing use from 72 hours to 96 hours. **American Journal of Infection Control. 26(1): 66-70**

RESEARCH DESIGN

Either the aim or objective and the type of quantitative approach was clearly stated within the abstract or summary of all the papers. Identification of the purpose and type of research at the beginning of the paper has several immediate advantages and limitations. From a nursing perspective, it is possible to determine if the research is relevant to patients in a particular care setting. However, as McCaughan (1999) points out, subtle differences may only be apparent with further exploration of the study. Secondly, clarity of purpose suggests a well-planned study that will deliver valid and useful data. Thompson (1999) concurs with this view and suggests that without a clear statement of aims the reader is unable to determine whether the research achieves it's objectives or not. Furthermore, with unclear aims, there could be a tendency towards data-trawling, which may result in providing spurious results. Although I agree with the views of Thompson (1999) and McCaughan (1999), I believe there can be concerns with stating

the type of method used in conducting the research. For instance, there could be a tendency to accept the design at face value or for it to bias the critical analysis of the study. Lai's (1998), study is described as a prospective, nonrandomised study. According to Woods and Catanzario (1988), a prospective design aims to observe a sample on at least two occasions over a period of time, the aim being to reduce the likelihood of bias in reporting the relationship between the cause and effect. Whereas Lai's (1998) study consisted of gathering information for a period of one month and could therefore be called prospective, it is not clear whether all the criteria for a true prospective design were applied or not. However, this is not to say that the research findings themselves would not be valid and reliable. Rather that even though there is a clear statement of the type of research, care must be taken that it does not influence personal judgement or the analysis of the study. Having determined that initially, the information given in the abstract of a study can have advantages and limitations, the next point will briefly consider the value of including or excluding a literature review within the study.

LITERATURE REVIEW

None of the studies reviewed in this assignment contained a literature review. However, there is some evidence that previous studies had been used to inform the investigations. Homer (1998) refers to previous studies within the methodology section and again in the discussion. The studies referred to are included in the reference list. Although the material in the reference list is relevant to the topic it only contains references for seven previous studies. Similarly, Lai (1998) cites previous studies within the discussion. Again the reference list only contains a few references. On the other hand, Curran et

al's (2000) study incorporates references to numerous previous work throughout the text. The value of including a literature review in a quantitative study is that it enables the research problem to be put into context or to identify gaps and weaknesses in prior studies so as to justify the current investigation. Also as Carnwell and Daly (2001) discuss it is valuable in demonstrating insight into the current state of knowledge within the relevant field. However, the exclusion of a literature review does not necessarily mean that the researchers did not undertake one. On the other hand, exclusion makes it difficult to determine the reason for undertaking the studies. For example, although, it appears that Lai's (1998) study may have been carried out, at least in part, to support a cost saving initiative within his medical centre, none of the referenced material concerns this issue.

However, there is a common understanding throughout the studies regarding the nature of the topic being investigated. In particular, all of the papers acknowledge that peripheral venous cannulae are associated with the development of phlebitis.

THEORETICAL/CONCEPTUAL FRAMEWORK

Although none of papers (Curran et al 2000, Homer 1998, Lai 1998) identify a theoretical or conceptual framework that guided the research, there is allusion to the principles of some of the theories. For example, Curran (2000) discusses two causes of phlebitis, insertion site infection and physiochemical reactions. Similarly, Lai (1998) refers to infusion-related phlebitis and sepsis. On the other hand, Homer does not appear to refer to any theoretical framework in his study. In conclusion, it does not seem that theoretical models had a major influence on the studies.

DEFINITIONS OF PHLEBITIS

The studies have all identified phlebitis by using clinical indicators and severity rating scales. However, there does not appear to be common indicators or scales to identify the severity of phlebitis. For example, Lai (1998) has defined phlebitis as a palpable cord or at least two of the following, tenderness, warmth, erythema and induration. There is no particular reference as to the origin of the definition, although it appears to have been adapted from the inflammation scale used by the IV team within the hospital where the study took place. Similarly, the origin of the inflammation scale is not reported. According to Lai's (1998) definition, phlebitis would be given a score of 3 or 4. Homer (1998) also defines phlebitis by clinical categorisation and allocating scores. However, the categories and scoring are different to those used by Lai (1998). Using Homer's (1998) definition would give phlebitis a score between 2 and 4. Homer's (1998) scale is not referenced but a significant difference from the one used by Lai (1998) is the exclusion of induration as a clinical indicator. Curran et al (2000) referred to the existing literature for indications of phlebitis. The eventual approach adopted contained two categories, erythema equal to or extending more than 3cm from the insertion site and/or purulent discharge with either or both being given a rating of 2 if present. Curran (2000) reports that this scale had been used in previous studies and includes references within the text, but does not offer any further information. However, there is no discussion as to any previous validation process that may have been used. As we have seen, phlebitis has been categorised and classified in different ways by different authors unique to their own study. Previous studies may have guided some of the

categorisation. For example, Homer (1998) refers to the classification of phlebitis by Maki and Ringer (1991 cited in Homer 1998 p5) and Tager(1983 cited in Homer 1998 p5) but develops his own scale for classifying phlebitis. As a result, there appears to be no consensus of opinion regarding the rating and definition of phlebitis between the studies under review.

As well as impacting on the accuracy of assessing and identifying severity of phlebitis, the lack of a common assessment tool for defining and measuring phlebitis may impact on the generalisability of the research. Campbell (1998) suggests that the use of a uniform scale that measures the degree of phlebitis is also beneficial in providing criteria for standardising documentation.

ACCURACY, VALIDITY AND RELIABILITY OF THE ASSESSMENT TOOL

Each of the studies used clinical indicators and a rating scale to determine and measure the severity of phlebitis. The strength and appropriateness of this type of measuring tool has been demonstrated in other quantitative studies within the same field. (Dinley 1976, Maddox and Rush 1977, Baxter Healthcare Ltd. 1988 cited in Campbell 1998 p 1311) However, there are limitations and weaknesses with the tools used in the studies under review. None of the researchers have discussed how validity and reliability was assessed. Although both Lai (1998) and Homer (1998) have used tools that do not appear to have been used in previous situations, there is no report on a pilot study being carried out prior to the main study. Lai's (1998) measurement tool has a number of limitations. Within the study Lai (1998) reports that the intravenous site was "monitored according to a well-defined

inflammation scale” and has published the indicators and rating for the inflammation scale in the study. However, the tool that Lai used to obtain the data on phlebitis in his study was not the reported inflammation scale. Lai (1998) appears to have used elements of the inflammation scale and designed a scale that has some differences from the inflammation scale that he reported was used. The inflammation scale uses pain and swelling as clinical indicators but in the scale that was used in the study swelling has been excluded and pain has reduced to tenderness. Another point of concern is the implication that the severity of phlebitis could be measured on a scale designed to measure the presence and severity of inflammation. This may lead to confusion between rating phlebitis and inflammation. In other words, it is not certain that the tool actually used measured what it was supposed to measure. Homer (1998) referred to other studies before adopting a different set of clinical indicators in his scale that he reports as being more rigorous than Maki and Ringer’s (1991 cited in Homer 1998 p5) scale but less so than Tager’s (1983 cited in Homer 1998 p5). However, there does not appear to be any reference to testing validity of the scale. A point of accuracy concerns the use of the scale for measuring infiltration as well as phlebitis. Although, the measuring of infiltration is not mentioned in the discussion section, the methods section gives the grading of infiltration as 1-3. The grading on the scale for phlebitis is 2-4. According to Homer (1998), the majority of the incidences of phlebitis were graded at 2. Consequently, phlebitis and infiltration were both graded as 2 using the same clinical indicators. Again, this raises the point of whether the scale measured what it was supposed to measure. The strength of Curran et al’s (2000) measurement tool lies in the fact that it had been used in other studies. However, there is no discussion as to the validity or reliability of the tool. It

is also unclear as to whether Curran et al (2000) carried out a pilot study or not. There is a reference in the paper to collecting data on forty catheters, which is not the total number of the whole data collection. However, although analysis was carried out on the data collected from this, there is no further discussion.

In summary, there appears to be serious weakness and limitations with the accuracy of the measuring tools. Validity and reliability appear not to have been tested. As McCaughan (1999) points out, if measuring tools are not valid then neither are study findings. Error due to the measuring technique used, the instrument itself or the person doing the measuring may affect reliability. Therefore, the next point will consider the issue of interrater reliability.

INTERRATER RELIABILITY

All of the studies used observation as the data collection method. Curran et al's (2000) study was multi-centred and involved thirty-nine data collectors. The study gives details of an instruction pack given to each collector. However, there is no discussion on the reliability level. As Woods (1988 cited in Woods and Catanzaro 1998 p 251) relates, interrater reliability is especially important when more than one rater makes repeated measures over time. Lai's (1998) study relied on an unspecified number of data collectors including himself, but interrater reliability is not mentioned. Homer's (1998) retrospective study does not mention who retrieved the data. Therefore it is impossible to ascertain whether it was carried out by more than one person. However, this issue is not discussed within the study. Ultimately, it is impossible to assess the reliability of raters in any of the studies.

SAMPLING

Lai (1998) and Homer (1998) have given clear indication of the type of setting for their sample. Curran et al (2000) has conducted a multi-centre approach that used surgical patients as the sample but there is no specific detail as to the type of centres included in the study. All of the studies used non-probability sampling. A weakness common to all of the studies is the lack of information regarding how the sample size was estimated. For instance, there is no indication that a power analysis was performed prior to the studies. According to Polit and Hungler (1995), a sample size that is too small may fail to collect clinically important effects or differences. However, Polit and Hungler (1995) also point out that when non-probability samples are used, a large sample may not reduce bias. Although there is a lack of information regarding inclusion and exclusion criteria for the sampling in all of the studies, Lai (1998) has briefly reported criteria for exclusion. The strength of Lai's (1998) prospective study lies in the apparent inclusion of all patients admitted during the month the research was carried out. Homer's (1998) study is particularly unclear with sampling methods. For instance, the retrospective study collected data on 722 patients over a period of three months. Firstly, this appears to be a relatively small sample for such a period of time. Secondly, it would appear that patients may have been excluded if documentation was incomplete. This would limit the number included and possibly limit generalisability. However, these points were not discussed in the study. Limitations with Curran et al's (2000) study concerns the lack of information regarding the type of centres used for the study. Although the sample was taken from surgical patients, there could be implications for generalisability. For example, if the majority of the centres

conducted minor surgical procedures as opposed to major surgery, the findings may not be representative of all surgical patients in general hospitals. Therefore, lack of information within all the studies inhibits the ability to be certain of the generalisability of the findings.

STATISTICAL METHODS AND DATA ANALYSIS

Each of the studies have used inferential methods to analyse the data collected. For example, from estimating risk parameters from the data collected, Homer (1998) has used a chi-squared test as a measure of association in order to test the hypotheses. However, the analysis becomes confusing with the use of risk models to further analyse the data. It is not possible to determine if the analysis is meaningful due to the complicated presentation of values. The concern to me was that the values are presented as mean plus or minus standard deviations. Firstly, without sophisticated manipulation of data, it would not be possible to obtain useful measures of standard deviation from the data collected as the distribution pattern was skewed. Secondly, the mean may not be a good measure of central tendency with an uneven distribution. For example, the range appears to be either sixty-three or twenty-four depending on whether the data collected was for infiltration or phlebitis together or phlebitis only. Consequently, I am aware that my lack of knowledge and familiarity with this type of test will need to be improved. Lai (1998) has used survival analysis to determine any percentage change in the number of lines that remained without evidence of phlebitis for the periods of seventy-two hours and ninety-six hours. Lai (1998) also uses Fishers Exact Test, which is an alternative to the chi-squared test for testing statistical significance between the two groups. As Polit and Hungler (1995) point out, Fishers Exact Test has the advantage that

it is suitable for small sample sizes. Curran et al (2000) has analysed data using various non-parametric tests that are appropriate to what was being tested. For instance, the Mann-Whitney U Test has been used to test the difference in the rank of scores of two independent groups whilst the Wilcoxon Signed Rank test was used to test the difference in the ranks of scores of two related groups. Overall the studies have given detailed accounts of the tests used for analysis of data. In contrast, Curran et al (2000) appears to be the only study that has taken account of any confounding variables.

DATA PRESENTATION AND FINDINGS.

All of the studies use tables or graphs to present descriptive and inferential statistics. These are accompanied by detailed written explanations of the findings. In particular, Lai's (1998) presentation is clear and it is possible to follow the analysis and verify the results. Homer's (1998) results are more difficult to follow due to the use of the risk model analysis. However, there is a full explanation of the results in the text. Curran et al (2000) has provided a table with the differences in variables between surveillance periods. However, the numbers in these do not always add up to the total number of catheters that were available for analysis. There does not appear to be any explanation for this. Furthermore, there is still lack of clarity with the sample numbers. Within the results discussion, Curran et al (2000) refer to the fact that the number of phlebitis episodes in surveillance period one varied from nought to ten out of forty or 0% to 25% and gives the median as two episodes. But in the table showing the differences in the variables, surveillance period one shows the phlebitis rate as being one hundred and

twenty five episodes out of one thousand four hundred and sixty three or 8.5%.

Ultimately, the findings from the different studies were very similar. Lai (1998) concludes that he is able to infer that the difference in the phlebitis rate between the two time intervals is not statistically significant ($p = 1.000$). However, he does acknowledge the limitations of the study and proposes that a randomised prospective study would be needed to confirm the findings. Homer (1998) reported his findings as indicating that re-starting therapy with a new cannula at seventy-two hours does not necessarily reduce the complication rate to the day one risk. The conclusion being that further studies are required to rationalise a policy of automatically restarting therapy after seventy-two hours. However, although Homer (1998) does not mention it in his results or discussion, the results of his study also suggest that infiltration may be more of a risk than phlebitis. Curran et al's (2000) findings suggest that the phlebitis rate for the second surveillance period was significantly lower than the first (5.3% v 8.5% p =less than 0.001). The conclusion drawn by Curran et al (2000) is that it is possible to be 95% confident that the reduction is probably between 20% and 50%. However, limitations with surveillance are discussed within the text. For instance, there may have been a Hawthorne effect. Overall Curran et al (2000) found four factors to be significantly associated with phlebitis. Whether an infusion pump was used, $p=0.002$, what the cannula was used for, $p=0.013$, duration of placement p = less than 0.0001 and the surveillance period $p=0.0007$. In respect of the length of time a cannula is in place, Curran et al (2000) recommend that there is no specific requirement to remove the cannula at seventy two hours providing there is no evidence of phlebitis/infection developing.

CONCLUSION AND IMPLICATION FOR PRACTICE.

As we have seen there are several threats to the internal validity of the studies reviewed. For instance, the extent of confidence that can be placed on both interrater reliability and on the assessment tool measuring what it was supposed to measure. Similarly, the lack of account of confounding variables in two of the studies (Lai 1998, Homer 1998) also suggests limitations with internal validity. Serious threats to internal validity may have an effect on the quality of the data collected and its compatibility with other relevant evidence. There are also some limitations with the sampling that may affect external validity. Lack of clarity in Curran et al's (2000) study regarding the setting and the sample numbers raised concerns over generalisability. Most importantly, this review highlighted that the lack of consensus over clinical indicators and rating of phlebitis is a potentially serious knowledge gap that has implications for risk management and safe practice. Another gap appears to be in relating the causes of phlebitis to theoretical or conceptual frameworks. Also, although it was not the prime rationale for the studies being undertaken, it has also shown that resource awareness and financial implications cannot be divorced from clinical practice.

In summary, all of the studies reviewed reached similar conclusions and recommendations concerning the length of time a peripheral venous cannula could safely remain in situ,

The studies reviewed in this assignment were three of fifteen retrieved from the literature search. The conclusions drawn from these three studies cannot to be used to inform best practice without further exploration of the relevant

literature. However, undertaking the review has highlighted the importance of critical appraisal of research studies before findings are applied to practice.

From a nursing perspective, the review has caused me to reflect on the current practice, assessment and documentation of what is considered to be a routine daily task, within my Trust.

In addition, conducting a critical appraisal of the studies for this review has highlighted areas for my own development. In particular, weaknesses with understanding some of the tests for inferential statistics. On the otherhand it has also raised my confidence in appraising research studies, a skill which I can continue to develop and share with colleagues.