Chapter IV:

Key findings

Chapter IV: key findings

4.1 Introduction

ASPIRE has aimed to assess the effectiveness of the Promoting Independence Programme (PIP) in Lower Hutt, Community FIRST (Flexible Integrated Restorative Support Team) in Hamilton and COSE (Coordination of Services for Elderly) in Christchurch. All are examples of ageing-in-place initiatives (AIPI), which specifically target older people with high and very high needs and support them to remain living in the community. The ASPIRE trial has sought to assess the impact of the AIPI on older people in preventing or delaying entry into permanent residential care as well as their role in reducing mortality. In addition, this project has aimed to evaluate the cost-effectiveness of the three initiatives and the results are available in ASPIRE (Report II). This chapter will present the key findings with some level of explanation surrounding each. The key findings are not presented in any order of priority.

The current Support Needs Level Assessment and categorisation system used by the NASC services to determine allocation of funding appeared to be highly variable across the three District Health Boards under investigation. It appeared that older people in Christchurch were assessed as being able to enter Residential care with a lower level of disability than those living in Hamilton and Lower Hutt. The interRAI MDS-HC assessment tool used by the research team provided a more rigorous and standardised method of assessment. This variation is probably a factor in variations between the key outcomes achieved in different services, such as reducing mortality or admission to residential care.

All three services appeared to reduce the risk of mortality compared with usual services. This varied from 28% in Community FIRST, 14% in PIP and 10% in COSE in comparison to older people in usual care. Although these figures are not statistically significant they reflect a clear trend in each service and are consistent with other results in the trial.

All three AIPI reduced the risk of entry to residential facilities when compared to usual care. Specifically, COSE reduced the risk of entry of older people to residential care in

comparison with the usual care NASC services by 43% (reduction). Community FIRST appeared to result in a reduction of risk of entry to residential care by 33%, in comparison to usual care, though given the lower sample size this was not statistically significant. The Masonic Promoting Independence Programme appeared to reduce risk of entry to residential care by 16% in comparison to usual care and as with Community FIRST; given the lower sample size this was not statistically significant.

Caregiver stress levels did not appear to increase in the intervention groups in comparison to usual care, despite the higher number of older people with high and complex needs remaining living at home.

An improvement in the independence levels of older people (Activities of Daily Living) within Community FIRST was noted, in comparison to usual care. No improvement in function was noted in the COSE or PIP initiatives in comparison to usual care.

Predictive modelling of the likelihood of older people being hospitalised or entering residential care was carried out using all the older people in the sample. This produced interesting results consistent with much overseas research.

- If a functional decline occurs in older people and the deterioration is not stopped, the older person is 11 times more likely to enter residential care.
- An older person is almost twice as likely to enter residential care if they are socially isolated.
- If an older person reports as having a negative mood, they are over twice as likely to be admitted to residential care.
- For every one unit increase on the Caregiver Reaction Assessment (which measures caregiver stress), there is a 7% increased risk of residential care entry.
- When an older person experiences inadequate meals and dehydration, they are over twice and 1.7 times more likely to be admitted to residential care, respectively.
- Delirium is highly correlated with risk of admission to residential care; those older people with delirium are 3.6 times more likely to be institutionalised.

- A lack of medication review (almost twice as likely), negative mood (1.5 times more likely) and previous hospitalisation (1.8 times more likely) are all correlated with increased risk of hospitalisation.
- The findings show there are a number of factors highly important to the enrolled population of older people such as coping, support, decision making and place of residence.

The study also explored the process by which older people entered residential care. Whilst the majority of older people often felt they had made the decision (to enter residential care) in most cases both the family and the NACS services thought that the family had been the main decision-makers. Further, nearly half of those who had entered residential care were sad or very sad about the decision. By contrast three quarters of those living in their own homes were happy or very happy with their decision to remain living at home.

4.2 Conclusions

Although the findings presented here can be viewed as is, it is important to note that interpretation must be undertaken in light of the cost-effectiveness of the relative ageing-in-place initiatives, which is available in ASPIRE (Report II). Also of note are the very different approaches each initiative took to facilitate ageing-in-place. Where as there are clear benefits in exploring multiple means to support older people to age-in-place, there is a tendency to compare the AIPI evaluated here and the relative success each achieved. In actuality, the strength of ASPIRE is to isolate those factors that are effective in facilitating ageing-in-place to allow new and existing services to evolve and develop.

Perhaps the most important finding of ASPIRE is that older people with high and complex needs who may enter permanent residential care are choosing on the whole to remain living at home with no apparent increase in risks to themselves.

Chapter V: Study Limitations

5.1 Introduction

Health service research is invariably challenging as it is not possible or indeed useful to create an artificial experimental environment. In order to generalise from the study, interventions need to be based in the real world and face every day common issues, such as recruitment and retention issues as well as referrals. The fact that ASPIRE was evaluating three services that were not developed or controlled by the research team provides considerable strengths but at the same time many limitations. This final section presents these limitations which must be considered when interpreting the results of this study.

5.2 Recruitment

The main study limitation is without doubt the low sample size in Hutt and in particular Hamilton. At the study onset, it was intended that recruitment would last 12 months with 12 months of follow-up. There were few issues in referrals, refusals or recruitment in Christchurch, but by 12 months, only approximately 100 were recruited to Hutt and less than 55 in Hamilton. At 12 months, a decision was made by the research team in collaboration with Waikato DHB and the MoH to continue recruitment in Hamilton for a further six months in an effort to increase the sample size. This naturally presents further problems as although the sample size increased past 100, almost 50% of the sample had follow-up data less than one year.

These difficulties in recruiting participants to ASPIRE in Hamilton presented particular challenges to the research team. Although the rate at which older people refused participation was not any higher than the other sites, the referrals from NASC were particularly low and required considerable intervention from both the DHB and the research team; as reflected by the considerably lower than anticipated sample size. On one level, this is perfectly understandable; Community FIRST had only been operating for a few months prior to commencement of the trial and therefore it would be anticipated that NASC

may have had concerns around the skills and abilities of staff within the new Community FIRST initiative. An important lesson from this trial for other DHBs considering adopting a similar approach would be to follow a mutually agreed implementation pathway.

The Masonic Promoting Independence Programme (PIP) had similar though not quite so dramatic issues in recruiting participants. As with Community FIRST, although the Masonic PIP had been running for some time in Levin, it was relatively new in the Hutt Valley and the service specifications required re-working prior and during implementation.

5.3 Design issues

When interpreting the relative effectiveness of all three services, it is important to consider the limitations of ASPIRE itself. ASPIRE provides a snapshot of how service are performing during a set time period and therefore if services are operating sub-optimally then the evaluation reflects that. All three services were undergoing a period of transition. COSE was beginning a roll-out from a very small pilot to almost half of Christchurch, Community FIRST had only been running for a few months prior to ASPIRE and were having issues in staff recruitment. The Masonic PIP experienced a change in service specification from slow stream rehab to PIP and further during that period experienced considerable staff turnover. Further, there was a loss of short term stay rehabilitation beds at the Masonic village with the preference to use other facility beds when required.

5.4 Selection bias

Selection bias occurs when an investigator attributes study results to the effect of the independent variable when in reality the results could be explained by differences in the subjects before the experimental intervention was implemented. Several issues arose over the course of the investigation that may have confounded data in this manner. The research team had little control over the referrals to the programme and therefore the sample may not be representative as health professionals may have decided that a number of potential participants should be relocated directly to residential care. Although there were many efforts to minimise this phenomena, there is a possibility that older people were pre-screened by the Multi-disciplinary team.

Randomisation facilitates an equal distribution of known and unknown confounders across groups and although statistical testing revealed that the AIPI and usual care groups were similar, there were some differences observed at baseline in the two groups such as memory loss which may have influenced the results over time.

5.5 Maturation

Changes that take place in subjects because of development or the passage of time rather than as a response to an intervention, can also pose a threat to internal validity. For example, analysis around independence levels (assessed by ADL) in Hamilton reveal a mean decline in the usual care group that may have accentuated the associated improvement in the LAY group. An inherent difficulty in researching older people is that there is an expectation that a number of participants will decline in function and health over the period of a lengthy study. Thus, the declines observed in the usual care group may be part of a normal age-related deterioration.

5.6 Testing

Assessments were undertaken multiple times on participants involved in the study. Participants may have remembered their responses to questions at previous assessment points are altered their answers as a consequence of being exposed to them several times.

A further difficulty arose in that, although all attempts were made to ensure that follow-up assessments were undertaken on the same time and day, on occasions the subject cancelled or re-scheduled appointments. The extent to which this occurred was minor, however, in a few cases, it was not possible to assess the subject at the same time and day and as a result, there is considerable potential for recording errors. To illustrate, 'morning stiffness' is a measurable phenomenon that seems to be associated with strength and other functional abilities, which appear to improve over the course of the day. Thus, although the number of occasions where this issue arose was few, there is always the possibility that data could be confounded in some manner.

Although the MDS-HC proved to be an invaluable means to draw comparisons regarding the independence and relative frailty of older people enrolled in each region, it also

appeared to have several key limitations. One of which was the perceived inability of the tool to detect change over time, particularly in relation to the ADL and IADL scale.

It is interesting that none of the services appeared to demonstrate a statistically significant change in IADL function, despite a statistically significant change in ADL function within Community FIRST. Therefore, there is a distinct possibility that this lack of change could be an artefact of the relatively poor scope of the IADL scale. When comparing the MDS-HC IADL scale with other scales assessing extended function, it omits important areas such as walking outside and crossing roads as well as a host of other items. Clinically, one would anticipate demonstrating a shift in IADL items before ADL items and therefore the lack of change observed here may be a product of the assessment tool, which appears to have been addressed in Version 3.0 release of MDS-HC.

Measuring quality of life is difficult and the EuroQoL 5D was utilised in this study more for from a cost utility perspective. Where as the full EuroQoL results are reported in ASPIRE (Report II), the visual analogue scale was presented in this report. There were no real clinically relevant trends observed across any of the sites in ASPIRE from either the older person's perspective on their quality of life or the caregivers. This is surprising, as one would anticipate observing improvements, if independence levels increased as observed in Hamilton. However, this could well be as a result of the considerably difficulty in assessing quality of life.

5.7 Hawthorne effect

The Hawthorne effect occurs when participants respond in a specific manner because they are aware that they are participants in a research project. This may have influenced the findings in some way, though it is impossible to determine to what extent.

5.8 Experimenter effects

Experimenter effects stem from characteristics of the researcher. Attributes such as age, gender and specifically profession and facial expressions may influence subject's behaviour and responses. To illustrate, when the subjects were being assessed they may have behaved differently because of their familiarity with the researcher. Indeed, several

researchers have pointed to subjects becoming so well acquainted with them that they exerted themselves in the assessments more than they would normally do so, in order to achieve better results. This issue is of particular relevance to the present study as many of the subjects had few weekly visits and would therefore look forward to the researcher attending for the assessment. Although these threats are minimised by the usual care group, it remains an inherent weakness.

Appendices

Appendices

Appendix 1: Statistical analysis

Additional sample size and cluster size calculations

The formula for Design Effect is:

DEFF = 1 + (m-1)* ICC

where m is the cluster size and ICC is Intra-cluster-correlation

As we can see from the formula above if the cluster size is large, even a small value of ICC can have a large effect on sample size. The sample size required for a cluster-randomised study is the sample size for individual randomised study times DEFF. Below are various scenarios of the DEFF in the Christchurch trial assuming 35% event rate in the control group.

Design Effect	Sample Size required in Christchurch	Total sample size required
1	277	832
1.5	416	971
2	555	1109
2.5	693	1248
3	832	1387

For the purpose of the setting the total sample size to allow 90% Power when all 3 centres are combined, we have assumed that the DEFF is equal to 2 (cluster size=11, ICC=0.1). This gives us the anticipated total sample size required, from the scenarios presented below, of 1109 patients.

Handling of Christchurch data

Since one of the study centres was a cluster-randomised trial, it needed to be handled differently from the other two individually randomised trials. Since the data within a cluster is correlated there was a need to consider clustering effect when conducting the analysis of the Christchurch data. Therefore, the standard analysis methods were not suitable for these data. This section discusses the analysis methods that account for the clustering effect.

Proc Phreg in SAS was used to handle dependent data by addressing the COVB (AGGREGATE) option in the procedure (Glidden, 2004; Aalen, 1995). Hence, the Hazard ratio adjusted with cluster effect will be presented. The next step was to utilise meta-analysis to pool the entire three hazard ratios together. Standard Meta-analysis does not allow for clustering effect. There are several statistical methods for the meta-analysis for cluster randomized trials, such as adjusted Mantel-Haenszel test, Ratio estimator procedure, woolf procedure, Generalized estimating equation approach, etc. However, in this instance, only the adjusted Mantel-Haenszel test was considered (Donner, 2001; Donner and Klar, 2002). Adjusted Mantel-Haenszel test is represented by the formula:

$$\chi^2_{CMH} = \frac{(\sum_{j=1}^{s}[A_{1j}(M_{2j} - A_{2j}) - A_{2j}(M_{1j} - A_{1j})]/[M_{1j}C_{2j} + M_{2j}C_{1j}])^2}{\sum_{j=1}^{s} M_{1j}M_{2j}A_j(M_j - A_j)/(M_{1j}C_{2j} + M_{2j}C_{1j} - 1)M_j^2}$$

where $M_{ij} = \sum_{l=1}^{n_{ij}} m_{ijl}$ is the number of subjects in intervention group *i* in trial *j amd* A_{ij} is defined as the number of successes in intervention group *i* in trial *j.* $C_{ij} = \sum_{l=1}^{n_{ij}} m_{ijl}[1 + (m_{ijl} - 1)\hat{\rho}]/M_{ij}$, which are the clustering correction factors (estimated design effect), where $\hat{\rho}$ is the estimated intra-cluster correlation coefficient. Then the 'clustered Mantel-Haenszel estimator' (Donner et al, 2001; Donner and Klar, 2002) is given by:

$$\hat{\Psi}_{CMH} = \sum_{j=1}^{s} W_{jC}\hat{\Psi}_j \Big/ \sum_{j=1}^{s} W_{jC}$$, where $W_{jC} = [\frac{C_{1j}}{M_{1j}} + \frac{C_{2j}}{M_{2j}}]^{-1}(1 - \hat{P}_{1j})\hat{P}_{2j}$

[To conduct analysis for secondary endpoints, a decision was made to ignore the clustering effect, when ICC is less than 0.05].

Appendix 2: The interRAI Minimum Data Set – Home Care (MDS-HC)

The MDS-HC has several inherent scales. These are:

Activities of Daily Living Scales

The MDS items have been combined to form two types of summary measures (Morris et al. 1999):

1. a single MDS ADL Self-Performance Hierarchy Scale and;
2. two versions of additive ADL scales based on the same item pool (MDS ADL Short Form and the MDS ADL-Long Form Scales).

ADL Self-Performance Hierarchy Scale

The ADL Hierarchy Scale is a measure of ADL performance and categorises ADLs according to stages at which they can no longer be performed. The aim of this scale is to reflect the disablement process rather than to simply sum reduction in function. The scale is based upon the four ADL items used in the ADL short form (i.e., personal hygiene, toilet use, locomotion, eating). Early loss ADLs (e.g., personal hygiene) are given lower scores than those lost at a later stage (e.g., eating). For each of these four items, potential difficulty is scored from 0 (independence) to 4 (total dependence). A 6-point hierarchical scale is created as follows:

Table 5-51: ADL self-performance hierarchy scale

Score	Description	Use of four ADL items
0	Independent	All four score 0
1	Supervision required	All four score 1 or less AND at least one scores 1
2	Limited impairment	All four score 2 or less AND at least one scores 2
3	Extensive assistance required (I)	Eating and locomotion both score less than 3 AND personal hygiene and toilet use both score 3 or greater
4	Extensive assistance required (II)	Eating OR Locomotion score 3
5	Dependent	Eating OR Locomotion score 4
6	Total dependence	All four score 4

These seven categories, which describe the older person's ADL self-performance, can be grouped into four impairment levels: relatively independent (score = 0 or 1), limited impairment (score = 2), extensive help (score = 3 or 4), and more severely or totally dependent (score 5 or 6). Morris et al. (1999) reported that the ADL hierarchy scale provides evaluations of ADLs similar to other established ADL measures and is also able to reliably assess change in ADL impairment levels over time.

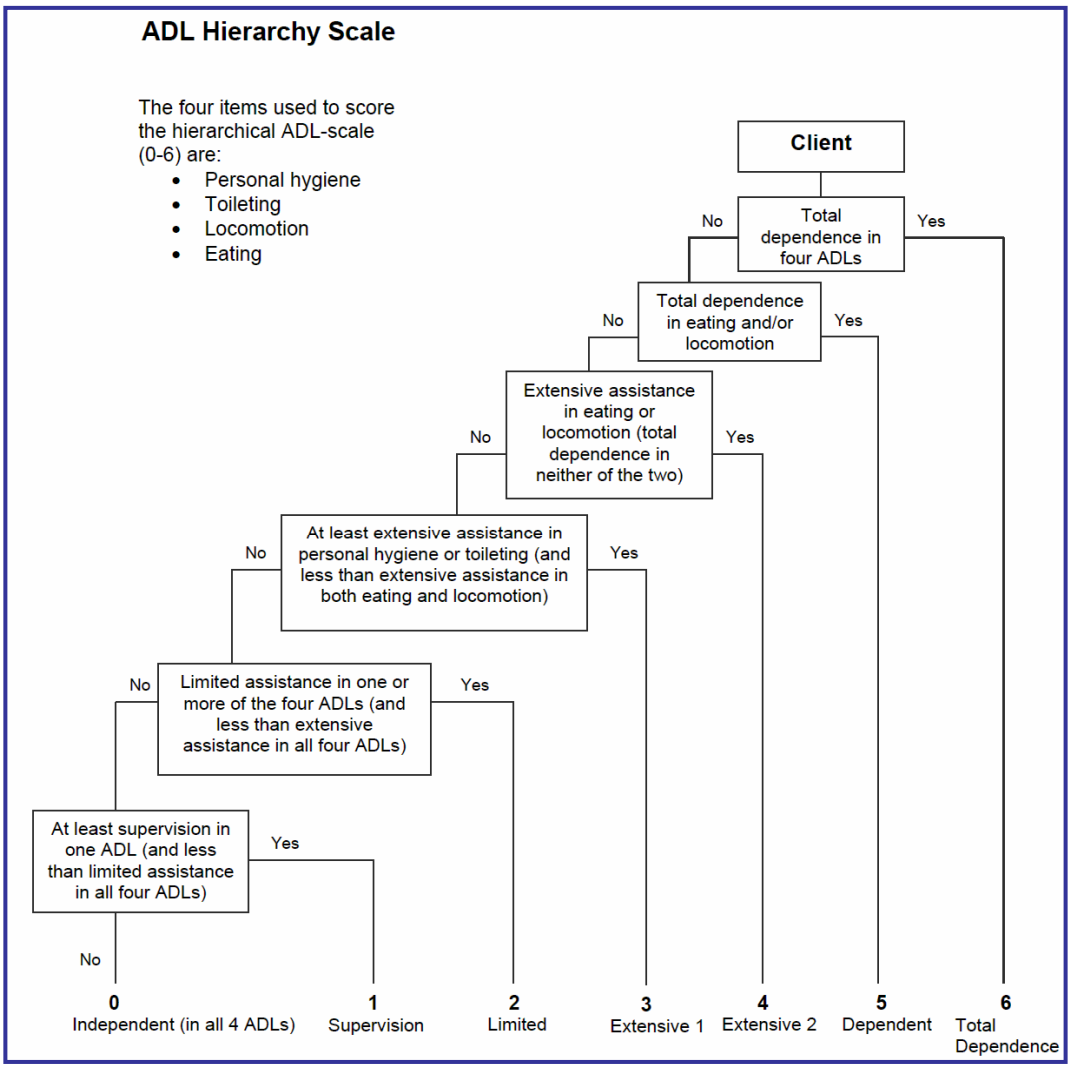

Figure 11: ADL Hierarchy Scale (Morris et al, 1999)

ADL Short Form Scale

This scale provides a summary measure of the client's ability to perform ADLs and is based on four items: Personal hygiene, toilet use, locomotion and eating. The scale has a range of 0 to 16, with higher values indicating greater difficulty in performing activities.

ADL Long Form Scale

The ADL Long form is a summary scale capturing all 7 of the ADL items. Each item is scored 0 to 4, creating a scale with a range from 0 to 28.

The Cognitive Performance Scale (CPS)

The cognitive performance scale (Morris, Fries et al. 1994) is a hierarchical index used to rate the cognitive status of older people. It has been validated against the Mini Mental State Examination, the Test for Severe Impairment (Morris, Fries et al. 1994; Hartmaier, Sloane et al. 1995) and includes four items: Short term memory, cognitive skills for daily decision making, expressive communication and eating. Based on the individual's impairment level on these four items, a CPS score scale ranging from 0-6 is derived (equivalent scores on the Mini Mental State Examination (MMSE) are shown below).

Table 5-52: CPS rating scale

CPS score	Description	Equivalent MMSE	Average
0	Intact	25	
1	Borderline intact	22	
2	Mild impairment	19	
3	Moderate impairment	15	
4	Moderate/severe impairment	7	
5	Severe impairment	5	
6	Very severe impairment	1	

Figure 5-13 highlights the scoring process in more detail.

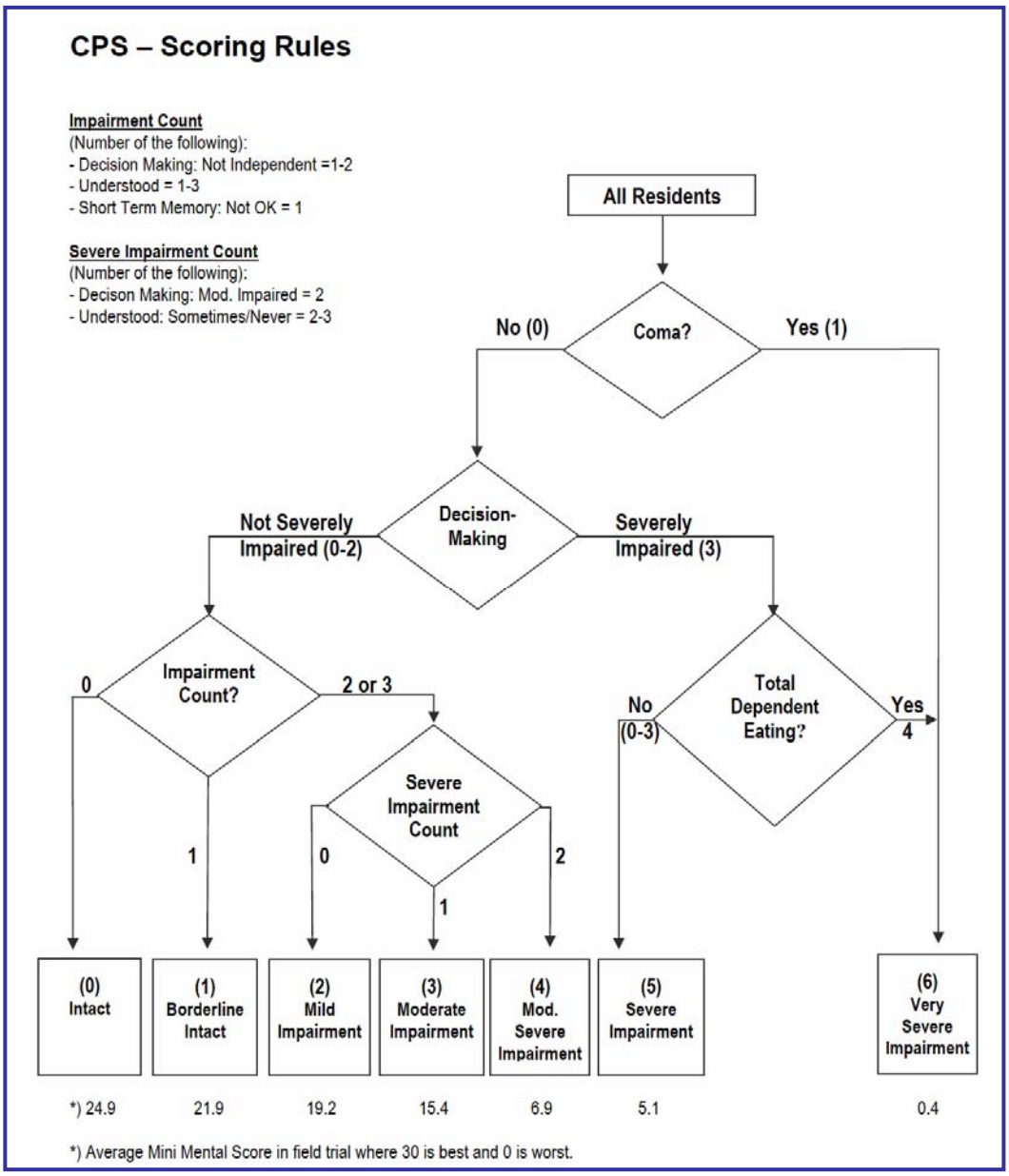

Figure 5-12: The Cognitive Performance Scale

Depression Rating Scale (DRS) (Burrows et al, 2000)

The Depression Scale can be used as a clinical screen for depression. It is based on 7 items embedded within the MDS-HC: Negative statements; persistent anger; expressions of unrealistic fears; repetitive health complaints; repetitive anxious complaints; sad, pained worried facial expressions and; tearfulness. Each of these 7 items is coded according to symptom frequency, resulting in a possible DRS score of 0 to 14. Burrows et al. (2000) demonstrated that scale scores of three or greater indicated major and minor depressive disorders in a nursing home population. They also established the criterion validity of this

scale through comparisons with both the Hamilton Depression Rating Scale and the Cornell Scale for Depression.

Table 5-53: The Depression Rating Scale (DRS)

Score	Indicators of depression
0, 1 or 2	Made negative statements
0, 1 or 2	Persistent anger with self or others
0, 1 or 2	Expressions (including non-verbal) of what appear to be unrealistic fears
0, 1 or 2	Repetitive health complaints
0, 1 or 2	Repetitive anxious complaints/concerns (non-health related)
0, 1 or 2	Sad, pained, worried facial expressions
0, 1 or 2	Crying, tearfulness

0 = Indicator not exhibited in last 30 days
1 = Indicator of this type exhibited up to five days a week
2 = Indicator of this type exhibited daily or almost daily (6, 7 days a week)

Instrumental activities of daily living (IADL) scales

IADL difficulty scale: The IADL Difficulty Scale is a hierarchical index that measures difficulty with three IADLs: ordinary housework, preparing meals and using the telephone. Client scores are combined to create a scale ranging from 0 to 6, such that higher scores reflect greater difficulty on these IADLs as follows:

Table 5-54: IADL difficulty scale

IADL Difficulty Score	Description
0	No difficulty on any of three IADLs
1-3	Some difficulty in 1 or more areas
4-6	Great difficulty in 1 or more areas

IADL involvement Scale

This scale is based upon a sum of 3 items in section H1 of the MDS-HC: ordinary housework, meal preparation and phone use. Individual items are summed to produce a scale that ranges from 0 to 9 (higher scores indicate greater difficulty in performing instrumental activities).

Table 5-55: IADL Involvement Scale

Score	Instrumental Activities of Daily Living
0, 1, 2 or 3	Meal preparation
0, 1, 2 or 3	Ordinary housework
0, 1, 2 or 3	Managing finance
0, 1, 2 or 3	Managing medications
0, 1, 2 or 3	Phone use
0, 1, 2 or 3	Shopping
0, 1, 2 or 3	Transportation

Scoring in self-performance:

0 = INDEPENDENT – did on own
1 = SOME HELP – help some of the time
2 = FULL HELP – performed with help all of the time
3 = BY OTHERS – performed by others

Changes in Health, End-stage disease and Signs and Symptoms (CHESS)

The CHESS scale was developed to detect frailty and instability in health. The CHESS attempts to identify individuals at risk of serious decline and can serve as an outcome where the objective is to minimize problems related to frailty (e.g., declines in function) in the elderly population. The CHESS scale was originally developed for use with the MDS 2.0 and has since been adapted for use with the MDS-HC. The CHESS first creates a subscale by counting across the following health symptoms:

Table 5-56: The CHESS scale

Assessment item	MDS-HC item code
Vomiting	K2e
Dehydration	L2c
Leaving food uneaten	L2b
Weight loss	L1a
Shortness of breath	K3e
Oedema	K3d

This score takes on values of either 0 (no symptoms), 1 (at least one symptom present) or 2 (2 or more symptoms present). This subscale is then added together with a score of 1 for each of the items on end stage disease (K8e), decline in cognition (B2b) and decline in ADL (H3) to result in a 6 point scale with scores ranging between 0 (meaning no instability) to 5 (for the highest level of instability). In the long-term care population, there is a clear differentiation of all six levels of CHESS scores, and higher levels are associated with a reduction in survival over time (Hirdes et al. 2003).

Pain Scale for the Minimum Data Set

The Pain Scale for the MDS was initially developed for use in nursing homes and later translated for use with the MDS-HC. The scale uses two items on the MDS (K4a and K4b) to create a score that ranges from 0 to 3 whereby 0 = no pain, 1 = less than daily pain, 2 = daily pain but not severe and 3 = severe daily pain. The Pain Scale has been shown to be highly predictive of pain on the Visual Analogue Scale in nursing home residents in the US (Fries, Simon et al. 2001).

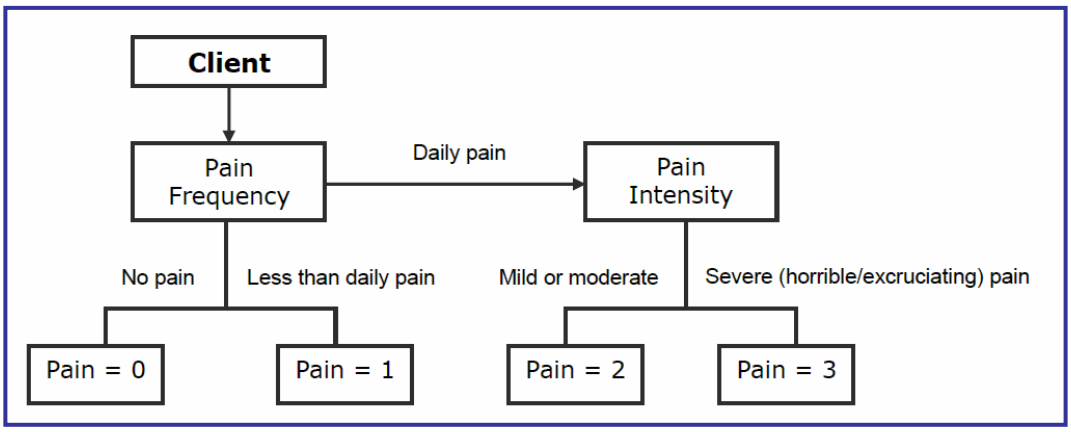

Figure 13: The Pain Scale

Appendix 3: Survival plots

The following figures are survival plots for firstly residential home placement and secondly mortality. The graphs provide a useful method of identifying the events (i.e. either permanent residential home placement or death) in the enrolled sample, as each drop in the plot indicates an event. However, events are relative to the enrolled population at that point and therefore the actual plots can be potentially misleading in small sample sizes as seen in Hamilton. For instance, if one older person is admitted to a residential facility at 18 months in Hamilton, it appears on the plot as being more of a decline than when one older person is admitted to a residential facility in Christchurch, as the enrolled sample in Christchurch at 18 months is larger and therefore one event may only be represented by a 5% drop (in a sample of 20), whereas in Hamilton it may be represented by a 20% drop (in a sample of 5). Notwithstanding this, the plots are useful up until 6 – 12 months into the study.

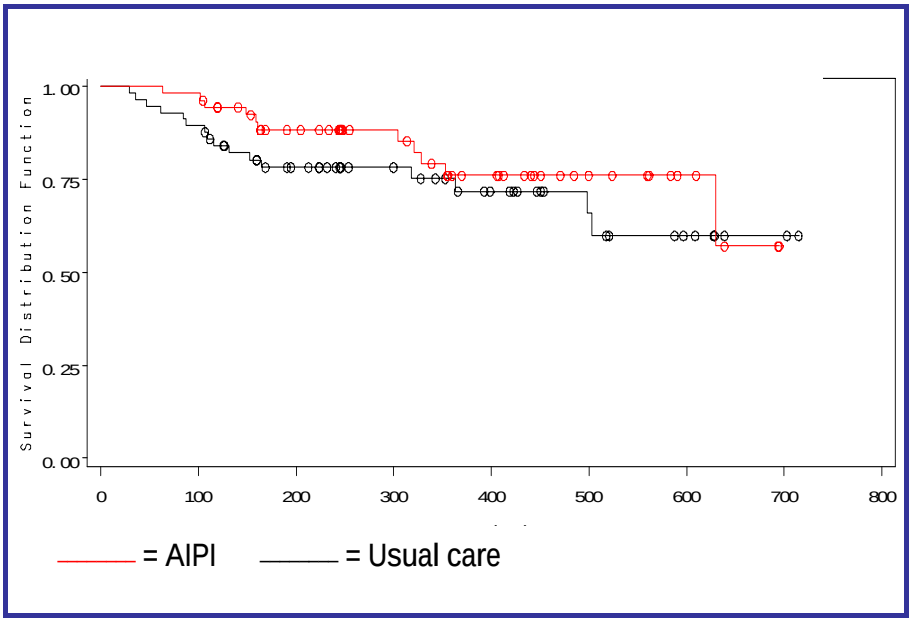

Figure 5-14: Kaplan-Meier curves between treatment groups in Hamilton region (using Death as Primary endpoint)

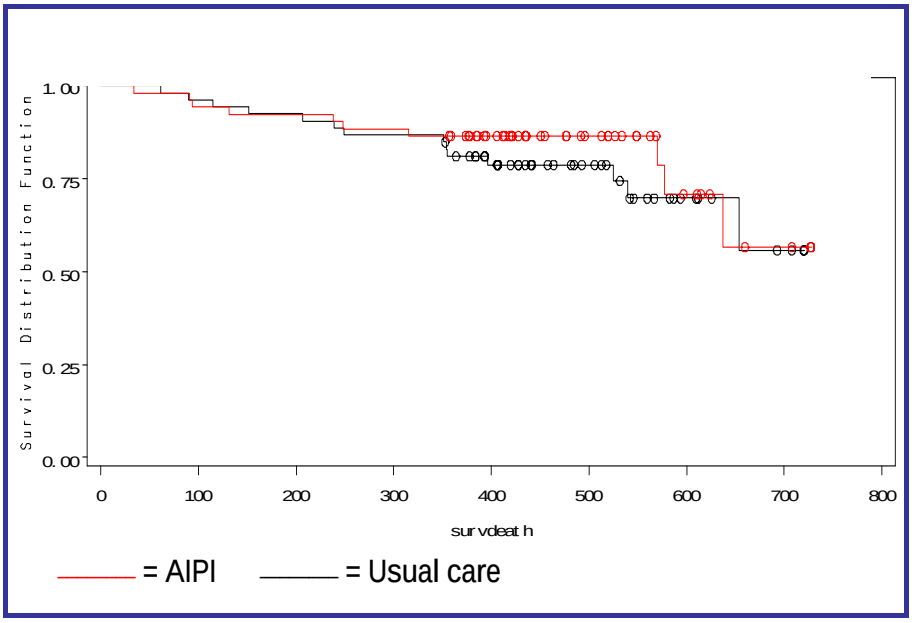

Figure 5-15: Kaplan-Meier curves between treatment groups in Lower Hutt region (using Death as Primary endpoint)

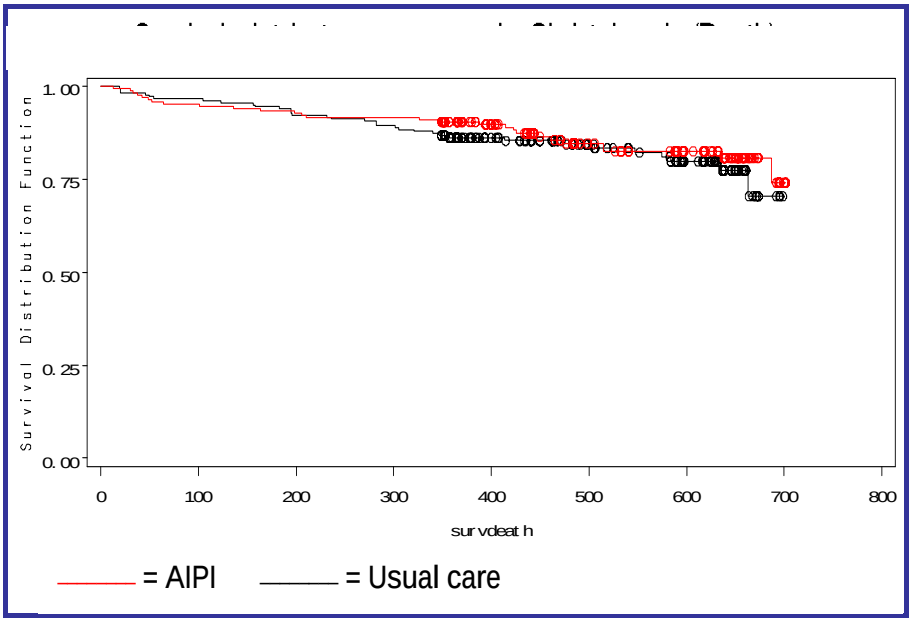

Figure 5-16: Kaplan-Meier curves between treatment groups in Christchurch region (using Death as Primary endpoint)

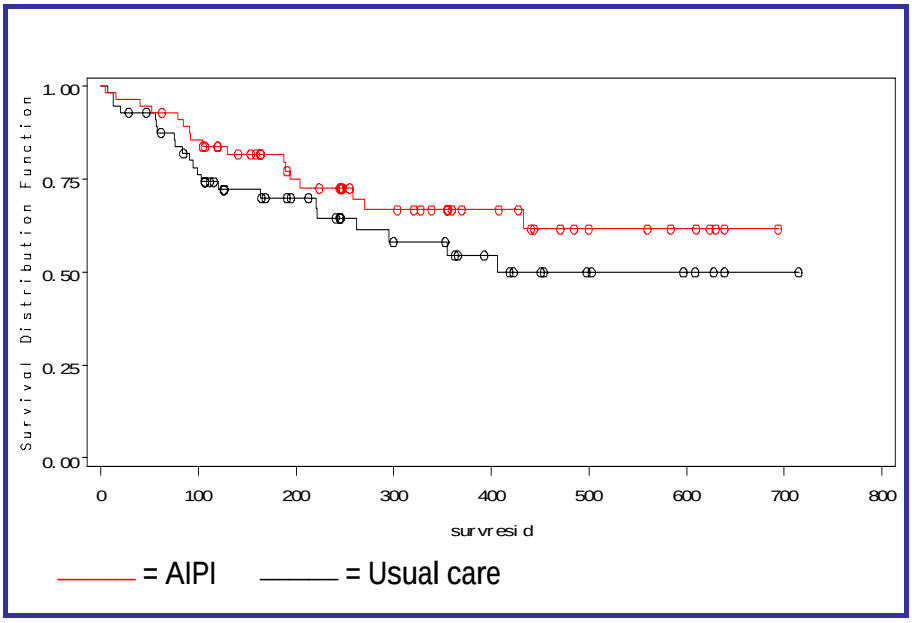

Figure 5-17: Kaplan-Meier curves between treatment groups in Hamilton region (using Residential Care entry as Primary endpoint)

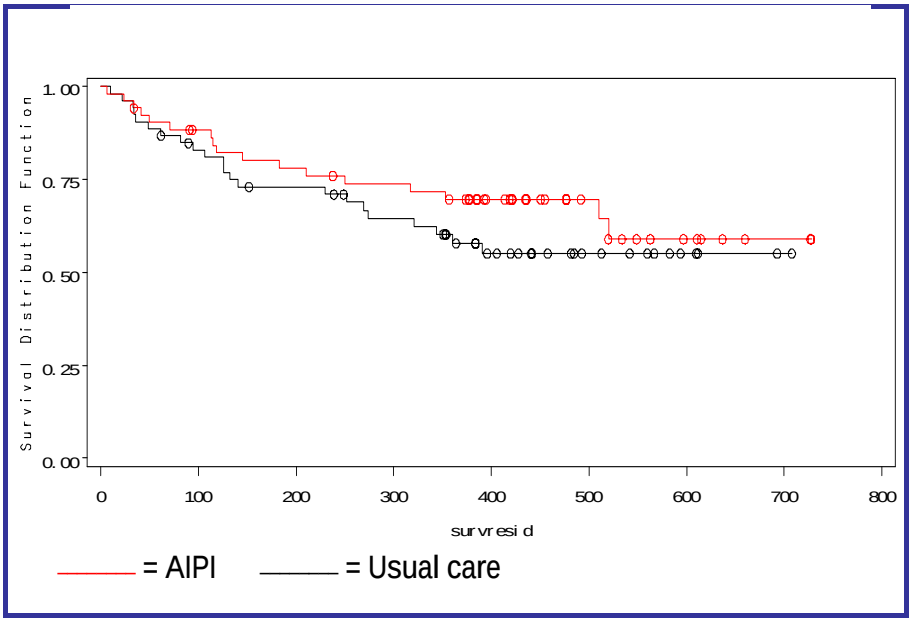

Figure 5-18: Kaplan-Meier curves between treatment groups in Lower Hutt region (using Residential Care entry as Primary endpoint)

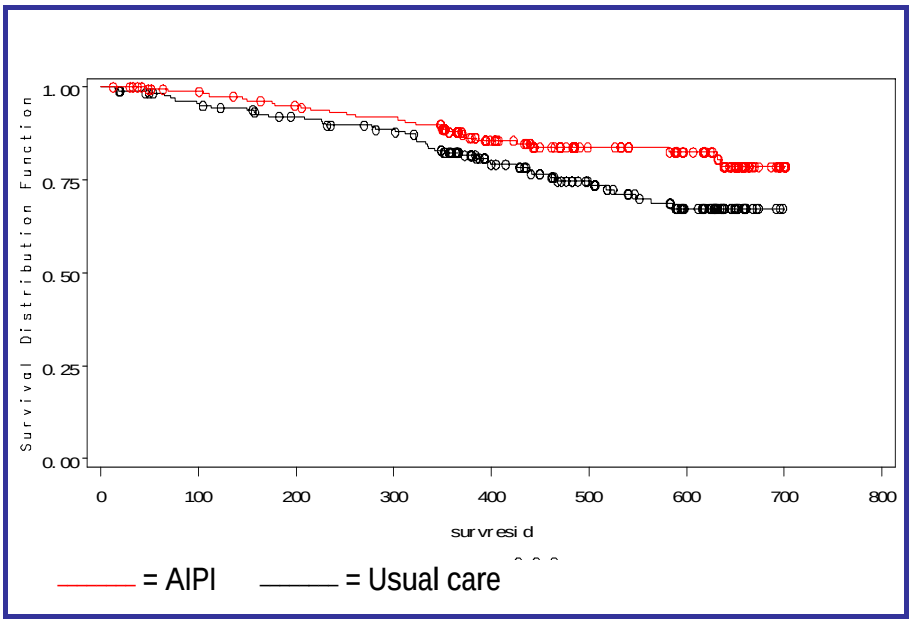

Figure 5-19: Kaplan-Meier curves between treatment groups in Christchurch region (using Residential Care entry as Primary endpoint)

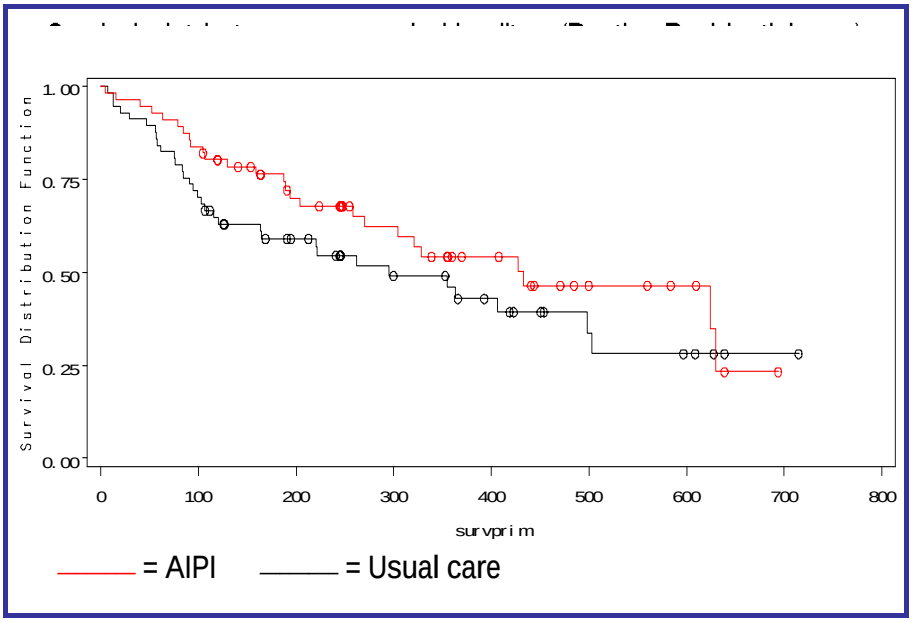

Figure 5-20: Kaplan-Meier curves between treatment groups in Hamilton region (using combined primary outcome as Primary endpoint)

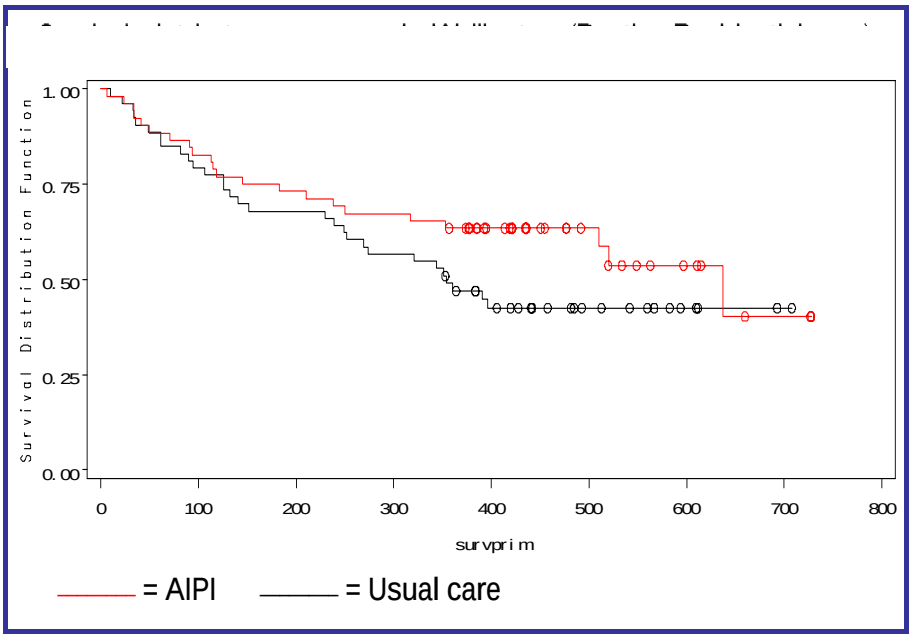

Figure 5-21: Kaplan-Meier curves between treatment groups in Lower Hutt region (using combined primary outcome as Primary endpoint)

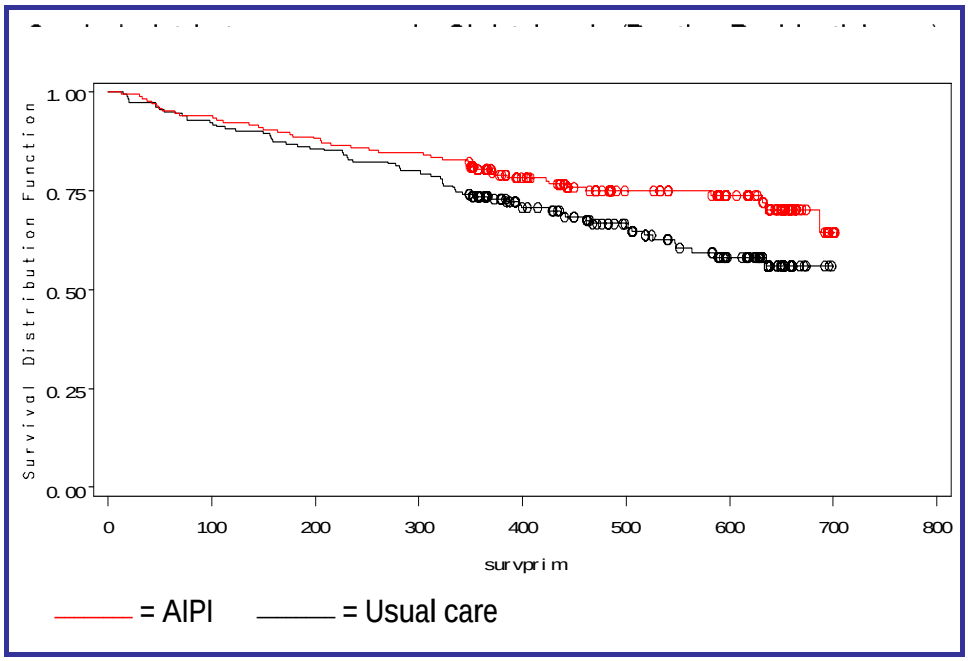

Figure 5-22: Kaplan-Meier curves between treatment groups in Christchurch region (using combined primary outcome as Primary endpoint)