



Review of the AR-DRG Classification Case Complexity Process: *Final Report*

1 August 2014

Prepared for the Independent Hospital Pricing Authority

Consortium Partners



NCCH
PO Box 170
Lidcombe NSW 1825
Australia

t: +61 2 9351 9772
f: +61 2 9351 9603
e: enquiries@accd.net.au
w: accd.net.au



This page intentionally left blank



Contents

Glossary of Abbreviations	ii
Executive Summary.....	1
1 Introduction	11
2 Data preparation.....	28
3 Using Adjacent DRG and diagnosis cost profiles to evaluate AR-DRG Version 7.0 Complications and Comorbidities Levels.....	30
4 The Episode Clinical Complexity Model.....	37
5 Treatment of the Principal diagnosis in classification design.....	68
6 Guiding principles for Diagnosis Complexity Level assignment	78
7 The use of condition onset flag in classification design	86
8 Evaluation of performance of the ECC Model	92
9 Continued refinement of ECC Model.....	112
10 Conclusion.....	120
11 References	124
12 Appendices.....	128
13 List of Figures	129
14 List of Tables	132



Glossary of Abbreviations

Abbreviation	Description
ABF	Activity Based Funding
ACCC	Australian Casemix Clinical Committee
ACCD	Australian Consortium for Classification Development
ACHI	Australian Classification for Health Interventions
ACS	Australian Coding Standards
ADRG	Adjacent Diagnosis Related Groups
ADx	Additional diagnoses
AIHW	Australian Institute of Health and Welfare
ALOS	Average Length of Stay
AN-DRG	Australian National Diagnosis Related Groups
APC	Admitted Patient Care
APR-DRG	All Patient Refined Diagnosis Related Groups
AR-DRG	Australian Refined Diagnosis Related Groups
CAC	Clinical Advisory Committee
CART	Classification and Regression Trees
CC	Complications and Comorbidities
CCAG	Classification Clinical Advisory Group
CCCG	Clinical Classification and Coding Group



Abbreviation	Description
CCL	Complication and Comorbidity Level
CDC	Coherent Diagnosis Class
CDs	Complex Diagnoses
CMG+	Casemix Group Plus
COF	Condition onset flag
DBC	Diagnose-Behandling Combinatie
DCL	Diagnosis Complexity Level
DRG	Diagnosis Related Groups
DTG	DRG Technical Group
ECC	Episode Clinical Complexity
ECCS	Episode Clinical Complexity Score
ECMO	Extracorporeal Membrane Oxygenation
ED	Emergency Department
GDRG	German Diagnosis Related Groups
GHM	Groupe Homogene Medical
HAC	Hospital Acquired Conditions
HAD	Hospital Acquired Diagnosis
HIM	Health Information Manager
HIS	Health Information Systems



Abbreviation	Description
HIV	Human Immunodeficiency Virus
HRG	Health Resource Groups
ICD	International Classification of Diseases
ICD-9-CM	International Classification of Diseases, 9th Revision, Clinical Modification
ICD-10	International Statistical Classification of Diseases and Related Health Problems, Tenth Revision
ICD-10-AM	International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, Australian Modification
ICD-10-AM/ACHI/ACS	International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, Australian Modification, the Australian Classification of Health Interventions and the Australian Coding Standards
ICU	Intensive Care Unit
IR-DRG	International Refined Diagnosis Related Groups
ICT	Information Communication and Technology
ITG	ICD Technical Group
JWP	Joint Working Party
LKF	Leistungsorientierte DiagnoseFallgruppen
LOS	Length of Stay
MDC	Major Diagnostic Category
MEDISGRPS	Medical Illness Severity Grouping System



Abbreviation	Description
MS-DRG	Medicare Severity Diagnosis Related Groups
NCCH	National Centre for Classification in Health
NHCDC	National Hospital Cost Data Collection
NMDS	National Minimum Dataset
NordDRG	Nordic Diagnosis Related Groups
OR	Operating Room
PA	Pricing Authority
PCA	Principal Clinical Advisor
PCCL	Patient Clinical Complexity Level
PDx	Principal diagnosis
RDRG	Refined Diagnosis Related Groups (Yale)
RID	Reduction in Deviance
SOI	Severity of Illness Index
UE	Unconditional Exclusion
UK	United Kingdom
US	United States
USA	United States of America
UWS	University of Western Sydney
V	Version



Executive Summary

Overview

This report presents the results of phase one of the project to develop Australian Refined Diagnosis Related Groups (AR-DRG)¹ Version (V) 8.0. This phase focuses on the case complexity component of the classification that has not been reviewed for many years.

Case complexity² was introduced as a feature of the Australian casemix system in the 1990s. The system in its current form was adopted in 2000 as part of AR-DRG V4.1. A review is therefore not only timely but essential given that the case complexity component of the classification has a significant impact on the measurement of a hospital's 'casemix value' which is the core component of Activity Based Funding (ABF) pricing models. To date, there have been multiple changes to the existing AR-DRG structure, however a systematic review of the structure's relationship to actual variations in resource use due to case complexity has not been undertaken.

Since 2000, patient data has improved in quality, especially in relation to diagnoses and interventions. As well, cost data is now available for the large majority of public hospital episodes, with hospitals increasingly using methods to directly cost individual episodes of care rather than approximate the costs through cost modelling methods. At the same time, the computing capacity to analyse large data sets has grown dramatically.

During the project, the Australian Consortium for Classification Development (ACCD) worked closely with both the DRG Technical Group (DTG) and Classifications Clinical Advisory Group (CCAG) in developing the proposed methodology for the Episode Clinical Complexity (ECC) Model. A DTG subgroup of clinical and classification experts from the ACCD and the DTG reviewed and formalised guiding principles for the scope of the Diagnosis Complexity Level (DCL) within the ECC Model (including unconditional and conditional code

¹ Throughout this report, unless otherwise specified, the term Diagnosis Related Groups (DRG) refers to AR-DRGs.

² Case complexity is a generic term which has been used throughout this report to replace terms such as 'severity' which can be confused with 'severity of illness', a different concept.



exclusions). For instance, codes which provide additional or supplementary information to another code already assigned have generally been excluded, with clearly explained exceptions.

Aims of the Review of the Case Complexity Process

The overall aim of the project as detailed in Item A.2.5 of the Schedule to the contract between the Independent Hospital Pricing Authority (IHPA) and the National Centre for Classification in Health (NCCCH) as lead of the ACCD was to:

‘...better explain the variation in costs occurring in the admitted patient data within the AR-DRG classification.’

This has been achieved through addressing IHPA’s requirements³ to:

1. Review the current Patient Clinical Complexity (PCCL) process and identify improvements and modifications.
2. Determine the codes considered significant (currently the Complication and Comorbidity (CC) codes) in measuring case complexity.
3. Determine whether there is a need for separate CC codes and/or matrix for paediatric and geriatric age splits.
4. Determine whether more levels of complexity for significant diagnoses are required (currently for medical DRGs there are three Complication and Comorbidity Level (CCL) values and for surgical DRGs there are four CCL values).
5. Examine whether more levels of complexity for the overall episode PCCL score are required (currently there is a maximum value of four).
6. Determine whether the condition onset flag (COF) data should impact the case complexity score when the COF value indicates that the condition arose during the current episode of care.
7. Validate codes that are to be significant to the DRG classification and the clinical reasonableness of the final case complexity results through clinical consultation.

³ Adapted from the Contract for AR-DRG Classification System Development and Refinement Services between IHPA and NCCCH (June, 2013)



Important terminology

To avoid confusion with the current case complexity system, a new terminology has been developed for use in describing **Episode Clinical Complexity** (ECC) which is the element of AR-DRGs that recognises and allows for cost variation within Adjacent DRGs (ADRGs). The following terms describe the various concepts within the proposed ECC Model.

- **Diagnosis Complexity Level** (DCL) is the case complexity weight assigned to each diagnosis within a particular ADRG.
- **Complex Diagnoses** (CDs) are the set (or list) of diagnoses that have a non-zero DCL within a particular ADRG.
- **Episode Clinical Complexity Score** (ECCS) is the measure of the cumulative effect of DCLs for a specific episode.

The ECC Model is based on DCLs that estimate relative costs associated with each diagnosis within each ADRG. As well, a new ECCS has been developed, to estimate relative costs associated with each episode. The ECCS is based on the DCLs of the diagnoses present in the episode, and is simpler than the existing PCCL, but continues the principle of giving the highest weight to the highest DCLs.

Methodology

An initial assessment was undertaken to determine whether the current system of measuring and classifying case complexity in AR-DRGs simply needed refinement or had to be redeveloped. This assessment involved consideration of the principles that should apply in constructing a case complexity process (including an examination of methods used in other countries), and a review of the explanatory power of the current system. It was concluded that a full review of the current structure was required as the current system lacked explanatory power and models used elsewhere could not be readily adapted to Australian conditions.

The next step of the process involved an exploration of factors that appeared to be associated with increased costs within ADRGs. The initial analyses of 2009-10 to 2011-12 patient and cost data were undertaken and two findings emerged. Firstly, that costs generally increased as the number of diagnoses increased. This was true whether or not the diagnosis was a principal diagnosis (PDx) for the ADRG, or not. Secondly, the degree of



variation in episode cost according to the PDx was noted. The PDx is currently ignored in estimating case complexity (except for neonates and obstetrics). The variation is substantial for surgical ADRGs, determined by the key intervention undertaken, but also noticeable in medical and other ADRGs. Consequently, the PDx was included in the method developed.

Consequently, the progressive change in cost for each specific diagnosis (including the PDx), as the number of diagnoses increases within each ADRG, was used as the base for the proposed case complexity method. An estimate of the relative cost related to a diagnosis in an ADRG has been defined by averaging these progressive changes.

Potentially there are a large number of relative costs defined in this way. Some classes of diagnoses have been excluded from the ECC Model, and their DCL is set to zero for all ADRGs. For many potential combinations of diagnosis and ADRG, there will be either no episodes with the combination or else a very small number. An algorithm is described for combining diagnoses and ADRGs to obtain a threshold number of episodes for statistically robust estimation, and to ensure comparability of DCLs across ADRGs.

To combine DCLs to estimate the ECCS, the ECC Model adds DCLs in descending order, using a decay component to adjust for the diminished contribution of multiple diagnoses vis-à-vis their individual contributions. The formula has a significantly simplified form compared to the existing PCCL formula, but is similar in application apart from a reduced decay component.

Key Findings and Recommendations

Key Finding 1

A literature review and consultative process revealed that detailed information on the formal (i.e. theoretical) development of diagnosis level (CCLs) and episode level (PCCLs) case complexity measures was lacking.

Key Finding 2

A further literature review of case complexity systems used in other DRG classifications internationally did not reveal an alternative system that could be readily adapted for use in Australia.



Key Finding 3

The current method of measuring case complexity, the CCLs, exhibits very little (if any) correlation with cost. This was based on an in depth review of the existing case complexity system using three years of patient level cost and activity data from 2009-10 to 2011-12.

Recommendation 1

Based on Key Findings 1 – 3, ACCD in consultation with the DTG and CCAG recommends that a new conceptually based, formally derived and data driven case complexity system be developed for AR-DRG Version 8.0 and future versions of the AR-DRG classification.

Key Finding 4

As a measurement of diagnosis complexity, the new conceptually based and formally defined DCLs were shown to exhibit significantly higher correlation with costs within ADRGs compared to CCLs.

Recommendation 2

Based on Key Finding 4, ACCD in consultation with the DTG and CCAG recommends that the DCL measure of diagnosis complexity be adopted as part of a new case complexity system.

Key Finding 5

Unlike the existing system, the PDx has been recognised to contain important information on complexity over and above its use in allocating an episode to an ADRG. This is especially true for surgical ADRGs. Many principal diagnoses are in fact combinations of health conditions, and include vital information on conditions that would otherwise be additional diagnoses.

Recommendation 3

Based on Key Finding 5, given the noted degree of variation in episode cost according to the PDx, ACCD in consultation with the DTG and CCAG recommends that the PDx be included in



the construction of DCLs, reflecting the information contained in many principal diagnoses on the complexity of a case within its assigned ADRG.

Key Finding 6

The list of diagnoses permitted to be assigned nonzero DCLs has been guided by principles that aim to characterise the scope of the ECC model in terms of diagnoses considered relevant for DRG classification purposes. The diagnoses identified as out of scope are called exclusions, some of which have been excluded unconditionally and others excluded conditionally (depending on other diagnoses present) based upon guiding principles for DCL assignment.

Recommendation 4

Based on Key Finding 6, ACCD in consultation with the DTG and CCAG recommends that the proposed guiding principles for DCL assignment and list of identified exclusions (unconditional and conditional) be adopted.

Key Finding 7

In considering the potential role of the condition onset flag (COF) within the classification, ACCD had difficulty in defining what a condition arising during the episode of care meant in terms of its preventability. It was determined that removing codes associated with conditions arising during the episode of care (COF = 1) from the complexity algorithm would reduce the capacity of the classification to explain true cost differences between DRGs. It would potentially alter incentives to treat patients with risks of complication.

Recommendation 5

Based on Key Finding 7, ACCD in consultation with the DTG and CCAG recommends that the COF should not be used to exclude diagnosis codes from the DRG development process.



Key Finding 8

The new ECCS was shown to be a much improved predictor of cost at the episode level when compared to the current PCCL. Overall, the ECCS was shown to have the potential to greatly increase performance of the AR-DRG classification.

Recommendation 6

Based on Key Finding 8, ACCD in consultation with the DTG and CCAG recommends that the ECCS measure be adopted to estimate clinical complexity at the episode level.

Recommendation 7

Based on Key Findings 1 – 8, ACCD in consultation with the DTG and CCAG recommends that the proposed ECC Model which has shown to be a much improved predictor of cost at the diagnosis and episode level be adopted as the new case complexity structure for AR-DRG Version 8.0 and future versions of the AR-DRG classification.

Key Finding 9

ECCS performance was evaluated on paediatric and geriatric episodes and compared to that of the current PCCL measure. When compared to the PCCL measure, ECCS showed a much improved ability to minimise bias in cost estimation within ADRGs among both cohorts (i.e. minimising over and under prediction of cost).

Recommendation 8

Based on Key Finding 9, ACCD recommends that separate approaches for paediatric and geriatric episodes are not required, given the improved performance of the ECC Model in explaining cost variations for paediatric and geriatric episodes.



Key Finding 10

Changes in clinical care and improvements in data quality over time were identified as necessitating the ongoing evaluation and review of the ECC Model to ensure it is best suited to its proposed role in the AR-DRG classification.

Recommendation 9

Based on Key Finding 10, ACCD in consultation with the DTG and CCAG recommends that an ongoing and systematic approach be taken to evaluate and refine the ECC Model as part of the broader AR-DRG refinement process.

Implications

Stability of the Episode Clinical Complexity Model

Testing of the ECC Model's stability with respect to changes in data has shown that care is required when updating the model over time (see the *Continued refinement of ECC Model* section below). It is anticipated that 2012-13 data will become available shortly to allow stability to be tested over this additional year.

Episode Clinical Complexity Model implementation

The existing case complexity system uses CCLs which take integer values between 0 and 4, and episode PCCLs which take integer values between 0 and 4. Within these boundaries, 768 (non-error) DRGs are defined. IHPA's requirements 4 and 5 (see above) require a review of these levels.

The ECC Model allows for DCLs for each diagnosis in each ADRG, to take integer values between 0 and 5, and an ECCS for each episode, to take a value between 0 and 31.25. This new approach allows for greater freedom and precision when splitting ADRGs into DRGs.



A set of principles to be followed in construction of DRGs has been considered and endorsed by the DTG and CCAG (provided separately to IHPA). Using these principles will result in a statistically robust and stable AR-DRG system going forward, and also presents a unique opportunity to review the use of non-clinical variables (e.g. length of stay) in DRG construction.

Changes to episode grouping

The move to a case complexity system that better explains cost variations due to episode clinical complexity will result in changes to the way in which episodes group within DRG hierarchies. Specifically, episodes with previously unrecognised complexity will change from low severity DRGs (e.g. those ending in 'C' and 'D') to higher severity DRGs (e.g. those ending in 'A' and 'B'). Conversely, episodes with overestimated complexity will change from high severity DRGs to lower severity DRGs.

These changes will lead to shifts in DRG cost weights and will change DRG composition at the hospital, network/district, state/territory and national levels.

Private Hospitals

No estimation has yet occurred on the implications of the ECC Model for private hospitals.

Education about the Episode Clinical Complexity Model and its implications

The ECC Model represents a significant change in the consideration of episode clinical complexity and its potential to simplify and improve the AR-DRG system.

A dedicated education program by ACCD in conjunction with IHPA should be considered in the period leading up to implementation of Version 8.0 on 1 July 2016 so that those affected have the opportunity to understand and become familiar with the new ECC Model and its strengths and implications.



Next steps

ACCD is required to propose a new AR-DRG version to IHPA by 31 October 2014. The immediate priority is to develop clinically relevant episode clinical complexity splits for each ADRG, to replace the existing case complexity system.

A method to efficiently obtain informed clinical advice on the validity of proposed splits through CCAG will be developed. This process will be undertaken having regard to the principles for construction of AR-DRGs agreed by DTG and CCAG for AR-DRG development, notably the need for stability and statistical robustness, as well as avoidance of inappropriate splitting variables.

With an ECC Model that better explains cost variations due to episode clinical complexity, it is possible that the AR-DRG classification can be simplified, for example, a reduced dependence on non-clinical variables such as LOS and a reduced need for Pre Major Diagnostic Category (MDC) DRGs within the classification. These possibilities will be explored in the next phase.

Proposals for DRG changes that have been received will be analysed during the development of the new version, noting that the new ECC Model will impact on many of these proposals.

It has been noted that changes in clinical care and improvements in data quality over time necessitate continuous review of the AR-DRG classification. Therefore, an ongoing and systematic approach will be taken to evaluate and refine the ECC Model as part of the broader AR-DRG refinement process within V8.0 and future versions of the AR-DRG classification.

‘Based on the outcome of this review, ACCD will incorporate the approved new ECC Model into the splitting phase for development of AR-DRG V8.0.’



1 Introduction

This section provides an introduction to the Australian Consortium for Classification Development (ACCD) and its overall objective related to the development and refinement of the Australian Refined Diagnosis Related Groups (AR-DRG) Classification System. An overall aim of the case complexity review project has been included with the requirements to achieve this.

The underpinning aspects of the project are also presented in this section. A literature review of the historical and international considerations in approaches to case complexity processing in Diagnosis Related Groups (DRG) development to date provided a base from which to begin; the governance and consultation arrangements provided the means to engage and consult.

The new terminology for case complexity processing in AR-DRGs is introduced along with a brief project overview.

The University of Sydney's National Centre for Classification in Health (NCCH), in collaboration with the University of Western Sydney (UWS) and KPMG, has been contracted by the Commonwealth of Australia as represented by the Independent Hospital Pricing Authority (IHPA) to maintain, develop and improve the AR-DRG Classification System for the Australian Health System.

The ACCD led by the NCCH has been established to comprehensively address the work program. The three consortium ACCD partners have the following responsibilities:

- The University of Sydney through the NCCH has overall responsibility and provides leadership for the project. The NCCH, with its associated experts, undertakes the development and refinement work on the AR-DRG Classification System, and is responsible for the governance arrangements, communication with stakeholders, education and publishing.
- UWS is responsible for information communication and technology (ICT) systems maintenance and development, including the user interface for submissions and queries, and the (internal) ICT platform to manage the processing of proposals and



updates to the AR-DRG Classification System. UWS also undertakes certification of grouping software and the testing of software tools available for clinical coding.

- KPMG provides expert advice on AR-DRG change proposals and analysis methodology, and facilitates public consultative work for AR-DRG classification issues and refinement.

ACCD's objective is:

***'To manage the AR-DRG and ICD-10-AM development processes in an ongoing capacity...[and]...to deliver the quality refinement of the System based on clinical practice, that builds on the previous investments in the System.'*⁴**

There are two major components of the AR-DRG Classification System, namely, the development and refinement of:

- The International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, Australian Modification, the Australian Classification of Health Interventions and the Australian Coding Standards (ICD-10-AM/ACHI/ACS), and
- The AR-DRG classification.

A significant and timely portion of refinement within the AR-DRG classification component of the work program was to review the AR-DRG case complexity process. This important first phase of work was completed on 30 June 2014 and has set the groundwork for the development of AR-DRG Version (V) 8.0 due for release on 1 July 2015 (anticipated implementation 1 July 2016) and future AR-DRG versions.

The overall aim of this project as detailed in Item A.2.5 of the Schedule to the contract between IHPA and the NCCH as lead of the ACCD was to:

'...better explain the variation in costs occurring in the admitted patient data within the AR-DRG classification.'

⁴ RFT IHPA 010/1213 AR-DRG Classification System Development Services



This has been achieved through addressing IHPA's requirements⁵ to:

1. *Review the current Patient Clinical Complexity Level (PCCL) process and identify improvements and modifications.*
2. *Determine the codes considered significant (currently the Complication and Comorbidity (CC) codes) in measuring case complexity.*
3. *Determine whether there is a need for separate CC codes and/or matrix for paediatric and geriatric age splits.*
4. *Determine whether more levels of complexity for significant diagnoses are required (currently for medical DRGs there are three Complication and Comorbidity Level (CCL) values and for surgical DRGs there are four CCL values).*
5. *Examine whether more levels of complexity for the overall episode PCCL score are required (currently there is a maximum value of four).*
6. *Determine whether the condition onset flag (COF) data should impact the case complexity score when the COF value indicates that the condition arose during the current episode of care.*
7. *Validate codes that are to be significant to the DRG classification and the clinical reasonableness of the final case complexity results through clinical consultation.*

'Based on the outcome of this review, ACCD will incorporate the approved new Episode Clinical Complexity Model into the splitting phase for development of AR-DRG V8.0 and future versions of the AR-DRG classification.'

This report addresses the above requirements and includes a literature review of the historical and international considerations in approaches to case complexity processing in DRG development to date.

⁵ Adapted from the Contract for AR-DRG Classification System Development and Refinement Services between IHPA and NCCH (June, 2013)



1.1 Background

Numerous classification systems are available to measure the severity of illness (Commonwealth Department of Health and Aged Care, 2000) and have been developed over time to allow for comparisons of hospital activity on the basis of some outcome (e.g. activity levels, costs, quality, etc.), once differences in the mix of cases have been taken into account. There are several definitions that might underlie casemix adjustment: severity of illness, risk of mortality, prognosis, treatment difficulty, urgency, and/or resource intensity of treatment (Averill, et al., 2008). DRGs and their international equivalents are generally concerned with grouping on the basis of the resource homogeneity for treating particular clinically-defined groups of patients.

DRG classification systems make use of a principal diagnosis (PDx) and additional diagnoses (ADx), previously referred to as secondary diagnoses, and a list of medical/surgical interventions determined at the point of clinical coding.

AR-DRGs were developed to reflect Australian clinical practice and use of hospital resources. All public and private hospitals in Australia code using ICD-10-AM/ACHI/ACS. The International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, Australian Modification (ICD-10-AM) and the Australian Classification of Health Interventions (ACHI) tabular lists include an annotation next to certain codes which indicate that an Australian Coding Standard (ACS) exists which will assist in the application of the code. ICD-10-AM and ACHI codes are subsequently grouped to AR-DRGs for acute admitted episodes of care. The AR-DRGs are used by public and private hospitals, and state and territory health authorities to provide better management, measurement and payment of high quality and efficient health care services.

The AR-DRGs classify units of hospital output. The classification groups acute admitted episodes into clinically coherent categories (outputs) that consume similar amounts of resources (inputs). An acute admitted episode can now be allocated to one of the 771 available AR-DRGs in V7.0. Almost half of the DRGs reflect differing levels of complexity within 406 broader groups known as Adjacent DRGs (or ADRGs).



All of the Australian DRG versions include a case complexity matrix. Each cell in the matrix represents the complexity added by a specific diagnosis within each ADRG, known to date as the CCL.

Given the elapse of time, a review of the case complexity system is timely. In some cases there have been significant changes in clinical practice (e.g. reduced length of stay (LOS)). Further, the availability of patient level data and associated cost information is much improved, and the computing capacity to analyse the available data is now far more superior than it was in the 1990s.

1.2 Brief history and development of DRGs in Australia

DRGs have a long history of development in Australia. In 1985 the first research in this area was undertaken to investigate whether the DRG classification system developed at Yale University in the United States of America (USA) was relevant to Australian clinical practice. The first release of the Australian National Diagnosis Related Groups (AN-DRG) classification occurred in July 1992. Initially, the USA editions of the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) codes were used as a basis for the development of the AN-DRG versions 1.0 (1992), 2.0 (1993) and 2.1 (1994). This was then followed by the use of the Australian First and Second Editions of ICD-9-CM as the building blocks for versions 3.0 (1995) and 3.1(1996) respectively.

Although not publicly released, AR-DRG V4.0 was a major update to AN-DRG V3.1. It was produced using ICD-9-CM Second Edition codes as an interim step in the move towards the introduction of ICD-10-AM. This version, which incorporated the use of the newly developedACHI (known then as Medicare Benefits Extended), provided the foundation necessary for ICD-10-AM/ACHI First Edition codes to be used as the base within AR-DRG V4.1.

Coinciding with the introduction of ICD-10-AM/ACHI/ACS in July 1998, AR-DRG V4.1 replaced the AN-DRGs in December 1998. This version required a repeated forward and backward historical mapping process from ICD-9-CM Second Edition to ICD-10-AM First Edition to ensure stability of the DRG grouped data. Logical maps were also used to increase the stability of the DRGs when undertaking this repeated mapping exercise. Using ICD-10-



AM/ACHI as a basis, the process of updating the AR-DRGs has generally occurred biennially to incorporate code changes made in each edition of ICD-10-AM/ACHI to date.

Over the life of the development of Australian DRGs a whole number change in the version indicates a major release (e.g. V4.0 to V5.0 and V5.0 to V6.0). A major release is where DRGs are inserted and/or deleted and a new structure introduced in the DRG classification. A minor update is indicated by a point increase in the version number (e.g. V5.1). Minor updates usually just incorporate new editions of the underlying ICD/ACHI classification system and an opportunity to address concerns in ICD code placement within the DRG groups.

1.3 Use of costing data in DRG Refinement

In the early 1990's classification development used length of stay (LOS) as a proxy for costs. Classification categories were created where different groups had different LOSs. Since the collection of the National Hospital Cost Data Collection (NHCDC), patient cost has been used to review DRGs in the development of AR-DRG versions. NHCDC coverage has progressively improved. There has always been a check used to ensure that LOS characteristics for hospitals supplying patient cost data were similar to LOS characteristics for the remaining hospitals.

Since the early work (1970's) on developing DRGs in the USA, changing clinical practice and improved technology has greatly reduced LOS for many conditions; and many interventions are now performed on a same day basis. Consequently, the usefulness of LOS as a proxy for cost in classification development has reduced. Furthermore, many of the differences seen in LOS are thought to relate to practice variations between hospitals rather than differences in the underlying care requirements of patients. For these reasons cost data is generally regarded as the best data available for classification development.⁶

⁶ Some countries make use of billing data that is, not what the patient costs the hospital to treat but how much the hospital charges for treating the patient. Where charges are based upon highly detailed itemised accounts, billing data might be a good proxy for actual cost, however the implicit assumption is that there is no cross subsidisation of different types of care within the billing process.



The usefulness of cost data for classification development depends upon the sophistication and appropriateness of each hospital's costing processes. In general cost modelled data are problematic as the patient groupings are used to allocate costs. In this case, analysis of cost differences may be confounded by the assumptions made in allocating the costs in the first instance. This means that the cost modelled data are not useful for classification development and so hospitals identified as cost modelled are not used in DRG and service weight development.

The Australian government has collected patient level cost information on patients treated in public hospitals for almost twenty years. These data are extracted by hospitals, sent to State and Territory health authorities who make modifications where necessary and then forward the data to the IHPA.⁷

There has been an ongoing effort to increase the quality and consistency of the NHCDC over time with the development of the Australian Costing Standards by the Commonwealth. These have been developed with reference to the United Kingdom Clinical Costing Standards, which were in turn informed by work undertaken by the Clinical Costing Standards Association of Australia (CCSAA). The CCSAA's functions are now undertaken by IHPA.

While there are ongoing efforts to improve the quality of the cost data, the quality of the data remain variable, depending in part on the extent to which individual hospitals have developed their costing processes, especially with respect to the following points.

- The extent to which hospitals rely on 'patient fractions' to differentiate costs across the different care types and the reliability of the 'patient fractions'.
- The accuracy and specificity to which hospitals map cost centres in their chart of accounts into the cost centre codes used for the NHCDC.
- The manner by which costs are allocated to individual patients (i.e. by actual costs, by cost weighted activity units, by bed day or hour, by service weight etc.).
- The experience of the costing staff and their understanding of the costing processes and national costing standard guidelines.

⁷ The Australian Department of Health and Ageing received the data prior to the establishment of IHPA.



Despite these on-going concerns with the quality of NHCDC data, the NHCDC by international standards represents one of the best and most comprehensive data collections of its type and is considered a sound basis for DRG development in Australia.

1.4 Current case complexity definitions

Treatment provided for a disease or condition can become difficult and more resource intensive and ultimately more expensive by the presence of a comorbidity or the development of a complication during the episode of care. AR-DRGs V7.0 identifies codes for CCs that are known clinically to contribute to higher resource consumption. The following terms are used within AR-DRG V7.0 with regard to case complexity:

- **Complication and/or Comorbidity (CC)** codes are diagnoses that are likely to result in significantly greater resource consumption.
- **Complication and Comorbidity Levels (CCLs)** are case complexity weights given to all diagnoses. The actual CCL value allocated to a diagnosis depends on whether the code is a valid CC, and whether it has been categorised as minor, moderate, severe or catastrophic in terms of the ADRG for that episode of care. Only CC codes are given a value greater than zero. Therefore a diagnosis given a CCL value of '0' is not a CC.
- **Patient Clinical Complexity Level (PCCL)** is a measure of the cumulative effect of a patient's CCs, and is calculated for each episode of care. The calculation is complex and has been designed to prevent similar conditions from being counted more than once.

1.5 Case complexity processing within the current AR-DRG classification⁸

The CC structure chosen as the basis of case complexity processing within AN-DRG versions was that used by the Refined DRG (RDRG) system, developed by Yale University in the late 1980's. The structure included a diagnosis exclusion list; to disregard diagnoses associated or related to another diagnosis already used to describe the case (some diagnoses provide additional information about a condition already coded and should not be considered a CC).

⁸ Section 1.5 is a brief summary of the work undertaken in the development of AR-DRG V4.0 which is further described in *Volume 3, AR-DRG Classification Version 4.0* (Commonwealth Department of Health and Aged Care, 2000).



To further develop and improve measurement of case complexity within AN-DRG V3.0, the Commonwealth undertook the Complication and Comorbidity Refinement project to examine the validity of the Yale CC structure.

The aim of the project was to specify CCLs appropriate for Australian clinical practice and to introduce new CCs based on recommendations made by the then Clinical Classification and Coding Group (CCCG). The project provided for Commonwealth ownership and specification of the AR-DRG CC list, CCL values and CC exclusion list.

The project involved a code level upgrading (recursive) process involving the CCL value suggested by statistical analyses based on the 1993-94 National Hospital Morbidity (Casemix) Database grouped to AN-DRG V3.0, and clinical consultation with various disciplines. Appropriate CCLs for those codes found statistically to increase the length of stay (LOS) and a corresponding CC exclusions list (3,215 codes) was then developed.

The Australian Casemix Clinical Committee (ACCC) Complications and Comorbidities Subcommittee recommended that the Commonwealth provide to the NCCH the current CC exclusions with descriptions and a list of the proposed additional CCs for review and drafting of a provisional set of CC exclusions for use with AR-DRG V4.0. While the ACCC's Complications and Comorbidities Subcommittee requested that all exclusion tables be reviewed by clinicians, this was not possible in the available time.

In addition, research was conducted on the effects of multiple CCs on resource use. In AN-DRG V3.0, a patient CC level is defined by the diagnosis with the highest CCL of all ADx within an episode of care. International research at the time however, determined that this approach did not adequately address the cumulative effect of significant ADx on resource consumption. An algorithm was therefore developed to create a new measure for PCCL. As a consequence, where there are multiple CCs, a patient may be assigned a higher PCCL value than a patient with only one CC.

Key Finding 1

A literature review and consultative process revealed that detailed information on the formal (i.e. theoretical) development of diagnosis level (CCLs) and episode level (PCCLs) case complexity measures was lacking.



1.6 Approaches to DRG development taken internationally to account for complications and comorbidities⁹

Identifying clinically homogeneous groups to better adjust for clinical complexity can be done in a variety of ways. Key differences are in how ADx are used to modify the base grouping (or ADRGs), whether and how weighting of ADx is undertaken, and whether and how any cumulative effects of multiple ADx are handled. These are described for major international systems below.

1.6.1 The United States

Initially, the Yale research team (Fetter et al., 1991) convened clinical panels to assess whether ADx codes with a particular PDx represented a substantial complication or comorbidity, using the joint criteria that it would add at least 1 day of stay for at least 75 per cent of patients affected (Averill, et al., 2008). Generic lists of complicating diagnoses were applied to most DRGs, but diagnoses closely related to the PDx were excluded for some DRGs. A set of 'unrelated' DRGs was created to account for patients whose in-hospital complications meant that they required surgery unrelated to the PDx, or reason for admission.

Use of DRGs for US Medicare hospital payments in 1983 gave rise to intense research efforts to refine definitions of complexity or severity, including the Severity of Illness Index (SOI) (Horn, et al., 1983; 1985); the Medical Illness Severity Grouping System (MEDISGRPS) (Brewster, et al., 1985), and Disease Staging (Gonnella, et al.1984), amongst others. Many required detailed chart review or additional data elements which were judged as not feasible (Gertman & Lowenstein, 1983; Smits, et al., 1983). A 1987 review of proposed refinements concluded 'there is no available measure of severity of illness that would

⁹ Section 1.6 is a summary of the discussion paper: *What approaches to DRG development have been taken internationally to account for complications and comorbidities?* Prepared by Associate Professor Terri Jackson with assistance by ACCD staff for purposes of the AR-DRG Classification System Development and Refinement project (see Appendix 1).



produce a large improvement in the accuracy of Medicare payments...' (Jencks & Dobson, 1987).

Outside the payment context, researchers also sought to predict how comorbidities might affect mortality outcomes. The Charlson Index (Charlson, et al., 1987) was the first of these, later expanded by Elixhauser, et al. (1998), and tested on other hospitalisation outcomes including acute admitted costs and LOS. Quan, et al. (2005) translated both algorithms into ICD-10. Both Charlson and Elixhauser (and their variants) identify a limited set of high-prevalence or high severity conditions, with weights derived from regression analysis on mortality or other outcomes. They have been used extensively as surrogates for acute admitted episode complexity, with and without DRG adjustment.

Recent UK research has shown that relative weights for these indices may have to be recalibrated for each hospital system because of coding conventions or system-specific clinical patterns (use of high cost drugs or imaging modalities, ICU capacity, etc.) (Bottle, & Aylin, 2011). Application to AR DRG-grouped data in Australia has shown the Charlson weights add little explanatory power (Jackson, et al., 2011).

Subsequent developments produced a hierarchical schema of ADx to modify DRG assignment with 3 or 4 case complexity levels in many DRGs. In 2008, with the introduction of Medicare Severity DRGs (MS DRGs), these complication and comorbidity lists were reduced to lessen the impact on the classification of stable chronic conditions, retaining only codes for 'significant acute manifestations' of such conditions (Averill, et al., 2008).

1.6.2 Germany

The German health care system adopted DRGs for hospital payment in 2003, basing their initial classification system on AR-DRGs. End classes were initially modelled on the Australian clinical complexity levels, but as local refinements were introduced, the initial 661 classes grew to 1137 by 2008 (Kalman & McCarthy, 2007). By 2011, the 1200 German DRGs (GDRGs) comprised 594 base DRGs with up to 9 complexity levels (Geissler, et al., 2011). More than two thirds of the end classes represent splits based on combinations of ADx and patient age.



1.6.3 France

France's Groupes Homogene Medical (GHM) classification development closely followed the US path, with simple without/with/severe CC splits. Version 11, however, entailed a major reconsideration of complexity adjustment, splitting most GHMs into four complexity levels and yielding some 2,300 end classes (Patris, et al., 2008).

In this revision, exclusion criteria were determined for groups of ICD codes, yielding 5 million pairs of ICD codes for exclusion. Lists of included ADx CCs were determined for each base GHM using iterative analysis of average LOS (ALOS) effects starting with the ADx with the greatest effect (the so-called 'isolated effect'), and successively adding ADx until no additional explanatory power remained. CC lists were then trimmed using LOS criteria.

1.6.4 Canada

Canada's classification development took a different direction, striving for maximum risk adjustment, regardless of the number of end classes. Their Casemix Group Plus (CMG+) system uses a range of factors weighted by regression coefficients (the extent to which each variable independently explains additional costs in a case) applied as adjusters to base casemix groups. In effect, coded patient separations are grouped into overlapping matrices by age cohorts, by 5 comorbidity levels, and by whether or not particular interventions (e.g. mechanical ventilation, return to theatre) were involved in their treatment. This results in a classification with thousands of end cells.

1.6.5 England

The Health Resource Groups (HRG) system is used in England. In terms of complexity adjustment of the base HRG, three levels are employed (without CC, intermediate CCs, and major CCs). Intervention complexity also drives HRG assignment, with 11 complexity 'bands' into which principal interventions can be assigned, and secondary interventions included in CC lists for some HRGs.



1.6.6 Other

Many countries have adopted or adapted casemix classifications from these early systems. NordDRGs, adapted for use in the Nordic and Baltic countries, are based on an early US system. Comorbidities are applied across the classification, rather than being specific to particular DRG groups. About 75 per cent of the DRG's appear in pairs that are as non-complicated (without CC) and complicated (with CC), i.e., the 'basic' DRG that includes both CC and non-CC cases is divided in two groups.

Other 'home grown' European classification systems provide few relevant examples of complexity adjustment. Austria's Leistungsorientierte DiagnoseFallgruppen (LKF) system is highly oriented towards interventions, and it has been noted that ADx or 'secondary diagnoses play a very minor role in defining case complexity levels' in the LKF system (Kobel & Pfeiffer, 2011). The PDx is not used (or checked) in assigning episodes to LKF for many interventions. Diagnose-Behandling Combinatie (DBC's) in the Netherlands characterise complex treatment pathways and include some forms of outpatient care. What other systems term 'secondary' diagnoses are considered to be co-equal primary diagnoses, potentially generating additional DBCs in a single episode.

It is notable that many countries use US developed casemix systems or the Australian DRGs, rather than undertake a separate development exercise.

1.6.7 In summary

Many health care systems have adopted acute admitted classification systems for a variety of purposes. All are influenced by the quality and depth of available clinical information in their record abstracts and by the coding classification system used to encode the clinical data. They have applied a range of methods to account for the wide variation in the treatment decisions of health care providers and in patients' basic physiology and response to treatment. There is little consensus on how increased complexity should be measured, whether it is generic or specific to each admission, and whether it should reflect only conditions present on admission, or other drivers of complexity such as the complications of acute admitted medical and surgical care.



The current PCCL measure was developed using constrained data and the concept itself has not been revised since its inception. Much of the earlier work on assessing CCs was based upon the extent to which LOS was increased, rather than costs. With the increased use of same day admissions the utility of LOS in describing cost differences is reducing

In light of the lack of consensus on how increased case complexity may be best measured, ACCD has developed an approach based on analysis of the sound statistical information on patient characteristics and costs now available in Australia.

Key Finding 2

A further literature review of case complexity systems used in other DRG classifications internationally did not reveal an alternative system that could be readily adapted for use in Australia.

1.7 Episode Clinical Complexity Model terminology

To avoid confusion with the current case complexity system, a new terminology has been adopted for use in AR-DRG V8.0 and future versions of the AR-DRG classification. The concepts associated with this new terminology are developed throughout this report.

Episode Clinical Complexity (ECC) is the element of AR-DRGs that recognises and allows for cost variation within ADRGs. The following terms describe the various concepts within the ECC Model

- **Complex Diagnoses (CDs)** in a particular ADRG are the set (or list) of diagnoses that have a non-zero Diagnosis Complexity Level within an ADRG.
- **Diagnosis Complexity Level (DCL)** is the case complexity weight assigned to each diagnosis within a particular ADRG.
- **Episode Clinical Complexity Score (ECCS)** is the measure of the cumulative effect of DCLs for a specific episode.



1.8 Governance and consultation process

The Pricing Authority (PA) has the overall governance role and is responsible for the proper and efficient performance of IHPA's functions. The final decision on the AR-DRG Classification System rests with the PA.

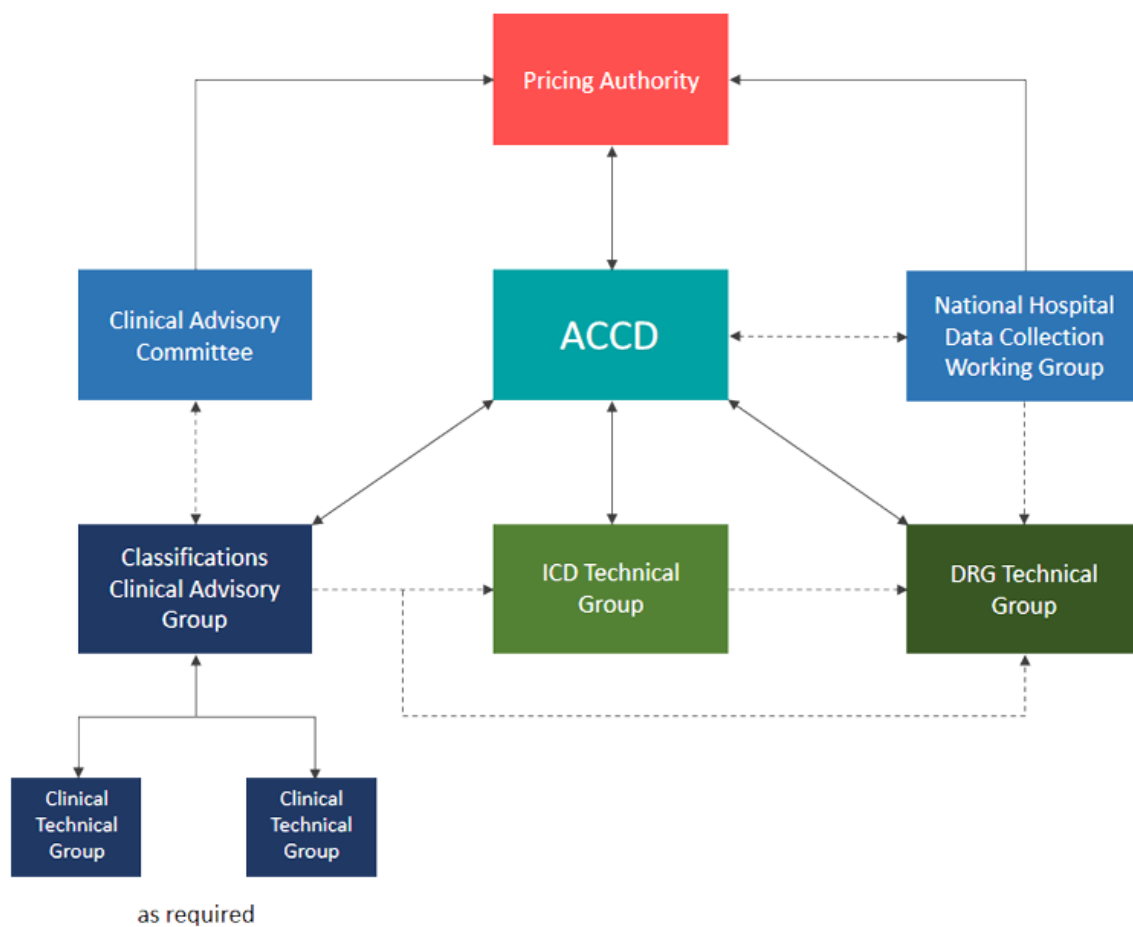
ACCD's governance arrangements, endorsed by IHPA include the establishment and management of the following technical groups to ensure appropriate communication channels:

- International Classification of Diseases (ICD) technical group (ITG): classification advice in regard to ICD-10-AM/ACHI/ACS.
- Diagnosis Related Groups (DRG) technical group (DTG): advice in regard to the refinement and development of AR-DRGs in Australia.
- Classifications Clinical Advisory Group (CCAG): to facilitate broad canvassing of clinicians to ensure that there is likely to be general acceptance of the developed proposals.
- Clinical technical groups: as required to provide specialty related clinical advice.

Figure 1 below depicts the AR-DRG Classification System Development and Refinement Services Governance Structure.



Figure 1: The AR-DRG Classification System Development and Refinement Services Governance Structure.



During the project, ACCD worked closely with both the DRG Technical Group (DTG) and Classifications Clinical Advisory Group (CCAG) in developing the proposed methodology for the ECC Model.

In this way, each Australian jurisdiction including the Commonwealth Department of Health and a number of key health organisations were exposed and had input into the development of the methodology for the ECC Model via their representative on the DTG. To date, the DTG met on four occasions, with the first full day face to face meeting held in February 2014.

Clinical input throughout the project was provided by CCAG. Since February 2014, CCAG met via teleconference on two occasions and had one face-to-face meeting on 23 June 2014. CCAG is chaired by a member of IHPA's Clinical Advisory Committee (CAC) and includes a



small number of clinicians with a broad knowledge range and interest/expertise in classifications with support from ACCD including the NCCH's Principal Clinical Advisor (PCA).

The initial analysis of the current case complexity system, the proposed ECC Model methodology, its continued progress and refinement and all associated discussion papers were progressively presented to the DTG and CCAG. The proposed approach was welcomed and the foundation for continued refinement along the lines of the proposed approach was endorsed by the DTG and CCAG.

A DTG subgroup was formed at the DTG meeting on 30 April 2014. Subgroup membership included clinical and classification experts from the ACCD and the DTG. The subgroup met on 27 May 2014 to review and formalise guiding principles for the scope of the DCL within the ECC Model. The outcomes of this subgroup meeting were presented and endorsed at the 3 June DTG meeting and CCAG on 23 June 2014.

1.9 Project Overview

This project forms Phase One of the overall development and refinement for AR-DRG V8.0 and future versions of the AR-DRGs.

The development of the ECC Model has involved a number of stages:

- *Data preparation* (page 28).
- *Using ADRG and diagnosis cost profiles to evaluate AR-DRG V7.0 Complications and Comorbidities Levels* (page 30).
- *The Episode Clinical Complexity Model* (page 37).
- *Treatment of the Principal diagnosis in classification design* (page 68).
- *Guiding principles for Diagnosis Complexity Level assignment* (page 79).
- *The use of the condition onset flag in classification design* (page 86).
- *Evaluation of performance of the Episode Clinical Complexity Model* (page 92).
- *Continued refinement of ECC Model* (page 112).
- *Conclusion* (page 120).



2 Data preparation

This section summarises the process used for data preparation, preceding the initial data analysis phase of the project.

2.1 Source data

The project draws upon data from six years of acute admitted episodes for public hospitals from 2006-07 to 2011-12, as reported in the Admitted Patient Care (APC) National Minimum Dataset (NMDS) and the National Hospital Cost Data Collection (NHCDC). The majority of modelling and analysis was undertaken using data from patient-costed public establishments reported in the 2009-10, 2010-11 and 2011-12 NHCDC. Supplementary analyses, including validation and sensitivity analysis, was undertaken using the earlier three years of the NHCDC, and some further analyses undertaken across the broader collection of APC episodes, including those from private establishments.

2.2 ICD-10-AM mapping

The diagnosis and intervention arrays of each year's APC data were mapped from the edition of ICD-10-AM native for that year forward to Eighth Edition. Morphology codes were removed from the diagnosis arrays during this process and episodes with diagnosis or intervention codes that were not valid with respect to the relevant edition of ICD-10-AM were identified and excluded.

2.3 Data exclusions

Changes to the classification of admitted patients in New South Wales and Victoria were back-cast onto earlier years to ensure consistency of definition of admitted patient episodes. Specifically, so-called emergency department (ED) only or 'ED-only' episodes, which were no longer classified as being admitted, were identified and excluded from earlier years of the data.



All episode costs reported in the NHCDC, except those reported within the depreciation cost bucket, were used. Costs were indexed to 2011-12 using a national indexation model that controls for changes in establishments and casemix (i.e. DRG and LOS).

Records were identified and excluded from the cost data on the basis of:

- being from establishments with less than 100 costed episodes;
- having identical costs across significant LOS intervals (indicating simple cost modelling);
- having implausibly low costs (\$25 or less); or
- advice received on costing anomalies.

Table 1 provides a summary of the final data used in the project. The table pools the cost data exclusions with the public establishment episodes without costs. The table excludes records removed through the ICD-10-AM mapping process and those identified as ED-only.

Table 1: Summary of final data used in the project (private and public establishments).

Year	Private (not costed)	Public (not costed)	Public (costed)	TOTAL
2006-07	\$2,826,152.00	\$2,419,095.00	\$1,911,673.00	\$7,156,920.00
2007-08	\$2,979,938.00	\$904,538.00	\$3,430,194.00	\$7,314,670.00
2008-09	\$3,084,132.00	\$835,388.00	\$3,626,264.00	\$7,545,784.00
2009-10	\$3,272,166.00	\$812,083.00	\$3,890,835.00	\$7,975,084.00
2010-11	\$3,354,398.00	\$712,905.00	\$4,134,248.00	\$8,201,551.00
2011-12	\$3,507,672.00	\$623,429.00	\$4,406,981.00	\$8,538,082.00



3 Using Adjacent DRG and diagnosis cost profiles to evaluate AR-DRG Version 7.0 Complications and Comorbidities Levels

This section details the initial analytical approach taken to evaluate the AR-DRG V7.0 Complications and Comorbidities Levels (CCLs).

The purpose of CCLs is to identify diagnoses that are associated with higher levels of resource utilisation and to quantify this higher level of resource utilisation relative to levels within each ADRG.

CCL values are integers ranging from 0 to 4; where 0 is intended to represent no associated higher relative resource utilisation and 4 is intended to represent the greatest relative impact. Medical ADRGs have CCL values ranging from 0 to 3, and surgical and other ADRGs have CCL values ranging from 0 to 4.

The review of the CCLs was undertaken by profiling the costs of ADRGs and diagnoses in such a way that allows comparative assessments to be made using associations between diagnoses and costs within each ADRG.

This section first introduces ADRG and diagnosis cost profiles, and then discusses the use of these cost profiles to evaluate CCLs within each ADRG.

Analysis was undertaken on the three years of costed episodes from 2009-10 to 2011-12. Particular diagnosis codes were excluded from the data prior to the analysis being undertaken; namely, external cause codes (U90.0 to Y91.9 and Y95 to Y98), place of occurrence codes (Y92.00 to Y92.9) and activity codes (U50.00 to U73.9).

3.1 ADRG cost profiles

ADRG cost profiles were derived by partitioning each ADRG into sets of episodes with common numbers of diagnoses. The diagnosis array within the data allows up to 100 diagnoses to be recorded against each episode, although generally there are significantly less than 100 diagnoses recorded against each episode, with less than 0.1 per cent of



episodes having 25 or more diagnoses (not counting the external cause and other code exclusions referred to above).

The following notation is used to describe the formation of ADRG cost profiles:

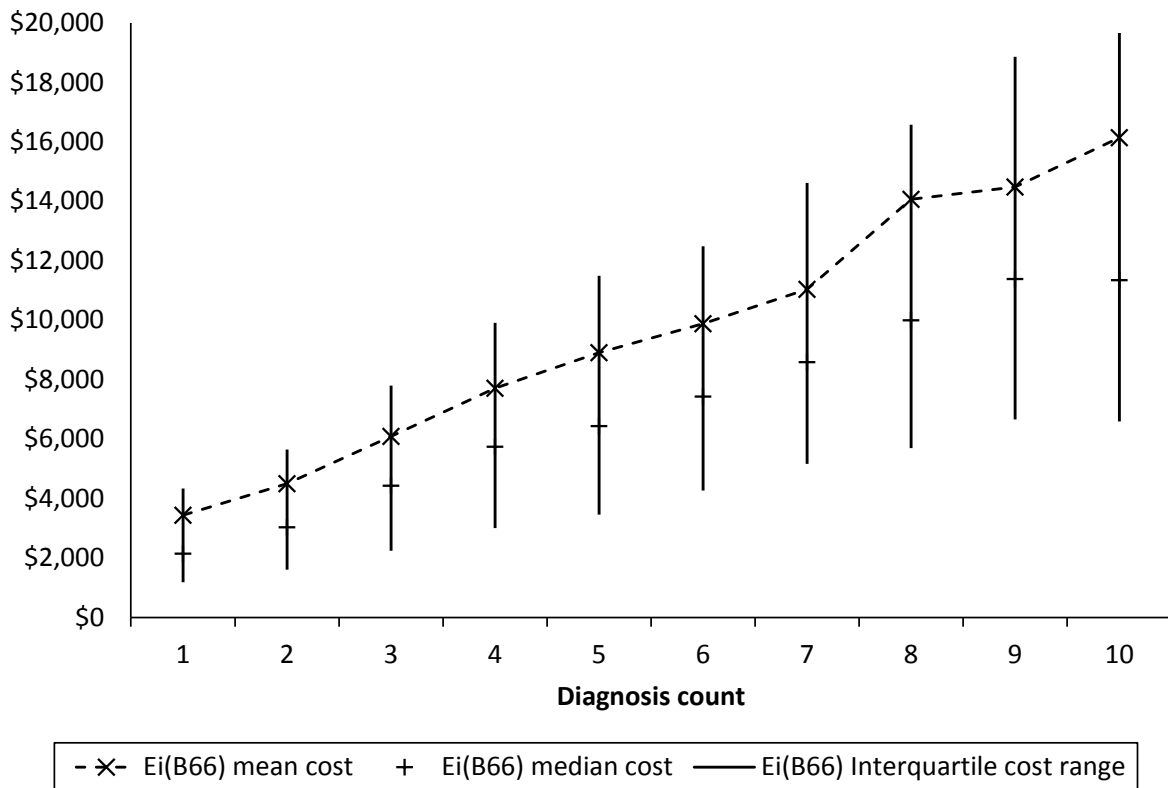
For each ADRG A and $i = 1, 2, 3, \dots$,

1. $E(A)$ denotes the set of all episodes that belong to A , and
2. $E_i(A)$ denotes the set of episodes in $E(A)$ with precisely i diagnoses.

Each $E(A)$ is the *disjoint union* of its $E_i(A)$ in the sense that every episode of $E(A)$ is contained in *one and only one* $E_i(A)$.

The episode costs of each $E_i(A)$ collectively form cost profiles that characterise the cost distribution of episodes with common numbers of diagnoses. These cost profiles can be examined individually or compared across i to examine the relationship between changes in costs and changes in diagnosis counts within each ADRG. Figure 2 illustrates how the $E_i(A)$ can be used to profile costs within each ADRG. The parenthesised numbers along the x axis of Figure 2 specify the count of episodes in each $E_i(B66)$.

Figure 2: Illustration of ADRG cost profiles – mean, median and interquartile range costs of $E_i(B66)$ for $i = 1, \dots, 10$.



3.2 Diagnosis cost profiles

The profiling of costs can be extended from the ADRG level down to the level of the diagnoses appearing within the episodes of each ADRG. The following notation is used to describe this process.

For each diagnosis x , ADRG A and $i = 1, 2, 3, \dots$,

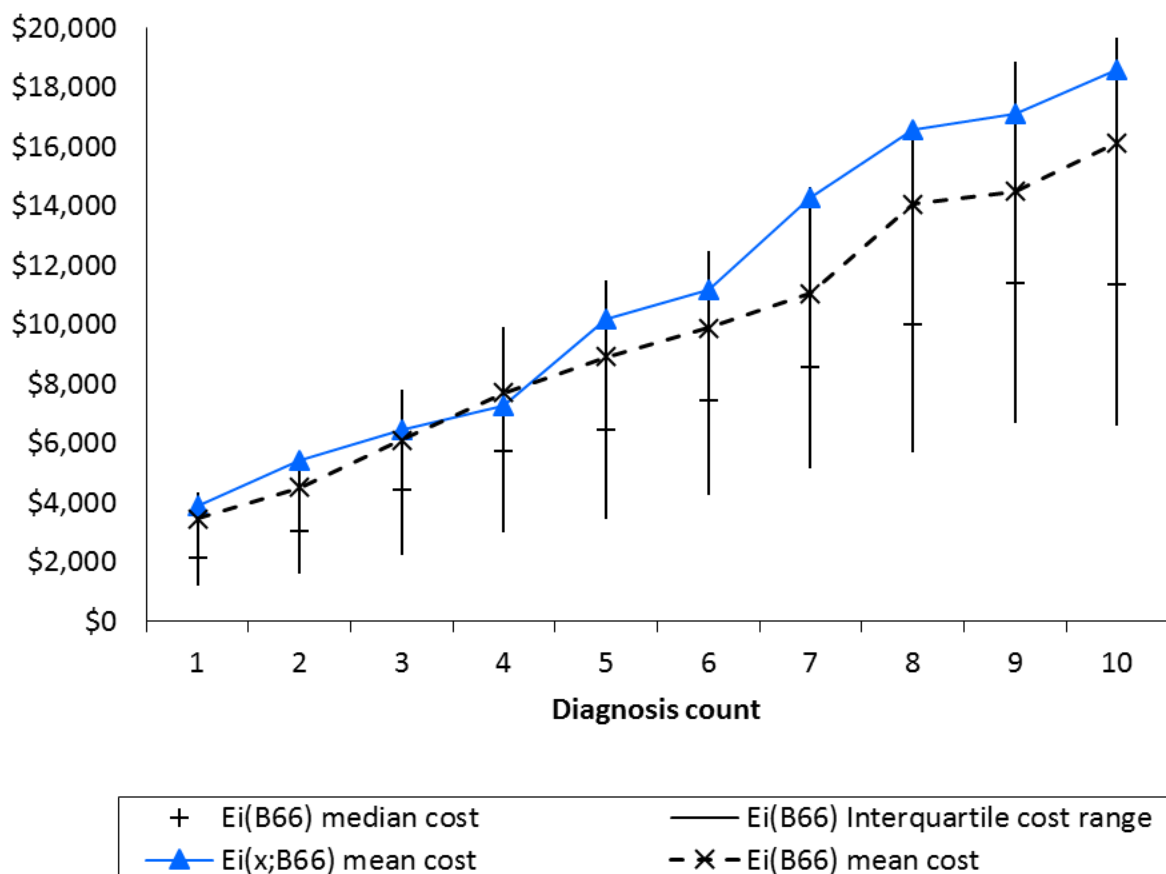
1. $E(x; A)$ denotes the set of all episodes in $E(A)$ that contain x , and
2. $E_i(x; A)$ denotes the set of all episodes in $E_i(A)$ that contain x .

It is important to note that the diagnosis x appearing in each episode of $E(x; A)$ or $E_i(x; A)$ can appear at any place in the diagnosis array. For example, an episode $e \in E_5(x; A)$ may have x appearing as its PDx (if possible) or as any one of its four ADx.

Similar to the relationship between $E(A)$ and its $E_i(A)$, each $E(x;A)$ is the disjoint union of its $E_i(x;A)$; that is, every episode of $E(A)$ containing a particular diagnosis x must be contained in one and only one of the $E_i(x;A)$. However, each episode within an $E_i(A)$ appears in multiple $E_i(x;A)$. For example, an episode $e \in E_5(A)$ with the five diagnoses x_1, x_2, \dots, x_5 appears in each of the five sets $E_5(x_1;A), E_5(x_2;A), \dots, E_5(x_5;A)$.

Each $E_i(x;A)$ has been used to profile the costs associated with x among episodes of $E_i(A)$. Figure 3 illustrates how the means costs of the profile of diagnosis x is comparable against the mean cost of $E_i(A)$ within a single $E_i(A)$ or across multiple $E_i(A)$.

Figure 3: Illustration of mean diagnosis profile costs of diagnosis x in B66.





3.3 Evaluation of AR-DRG V7.0 CCLs using diagnosis cost profiles

The effectiveness of the PCCL process in measuring patient clinical complexity is critically dependent on the ability of CCLs to quantify relative levels of resource utilisation associated with diagnoses within the context of ADRGs. Diagnosis cost profiles were used to evaluate CCL performance by measuring the correlation between diagnosis costs and their CCLs within each $E_i(A)$. Figure 3 and Figure 4 illustrate how this was achieved; showing scatter plots of diagnosis profile mean costs within $E_2(F14)$ and $E_3(F14)$ against each diagnosis' CCL value.

Figure 4: Scatter plot of diagnosis profile mean cost by CCL value for diseases occurring in $E_2(F14)$.

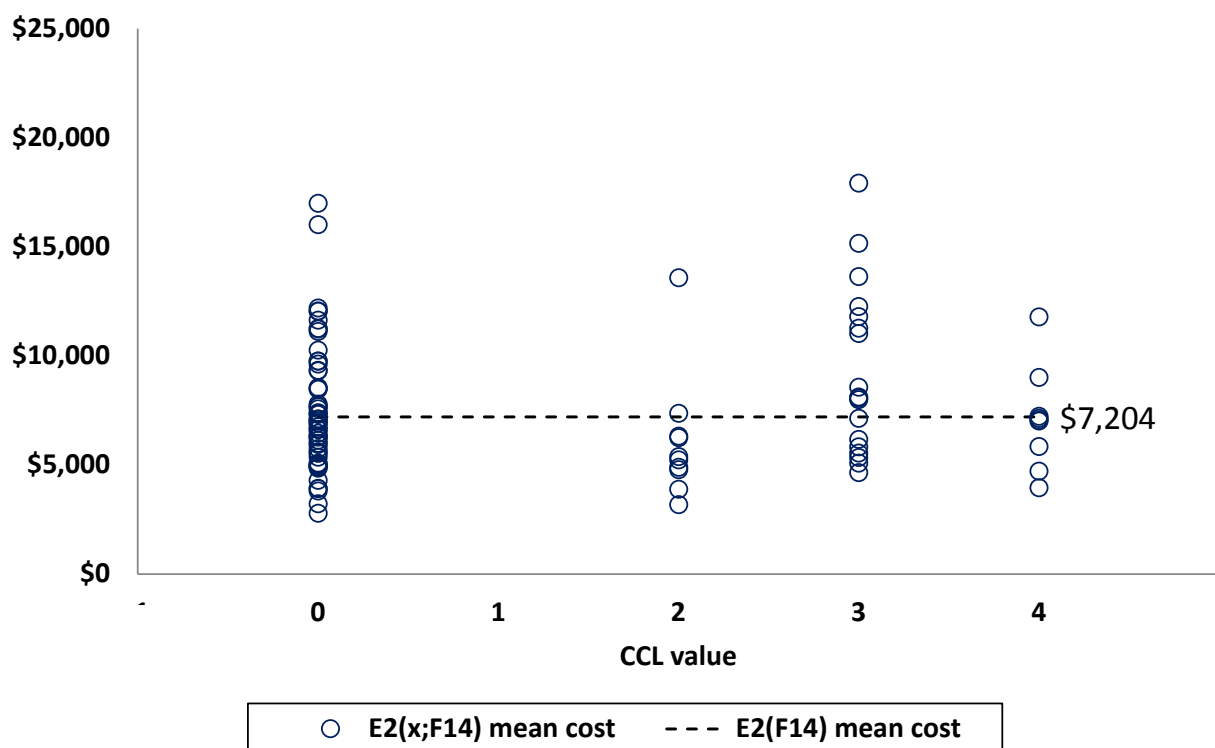
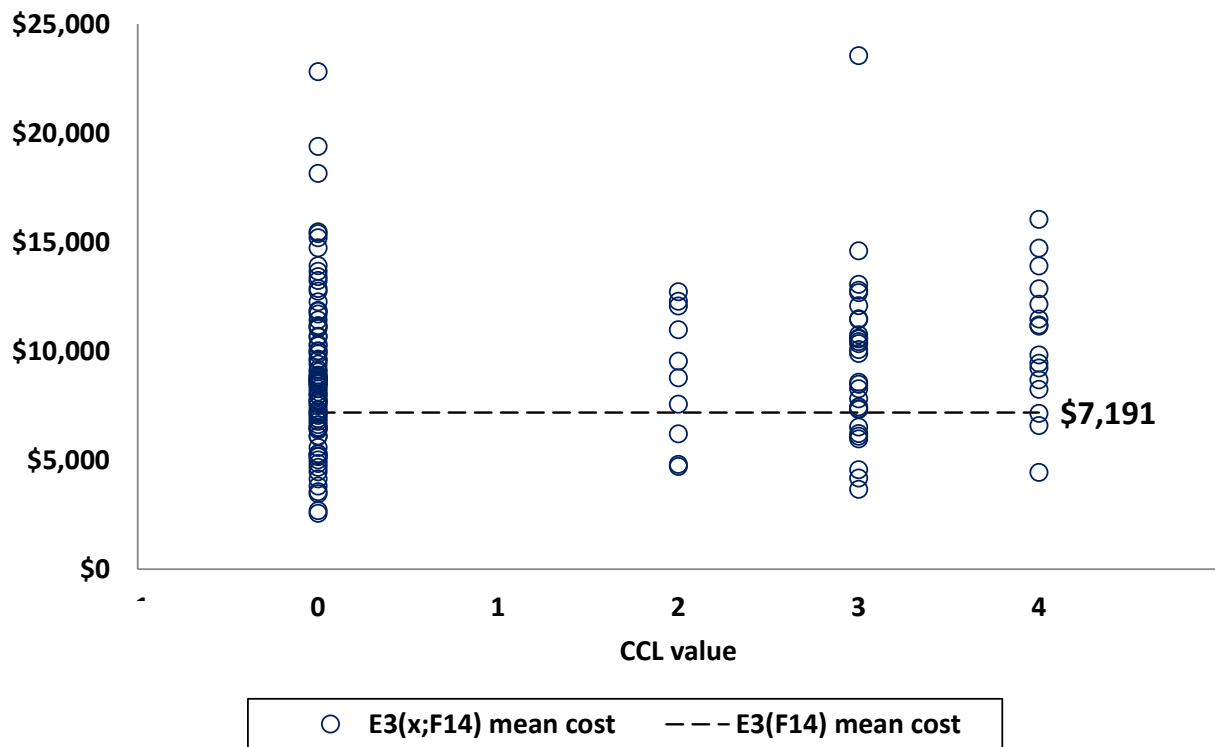




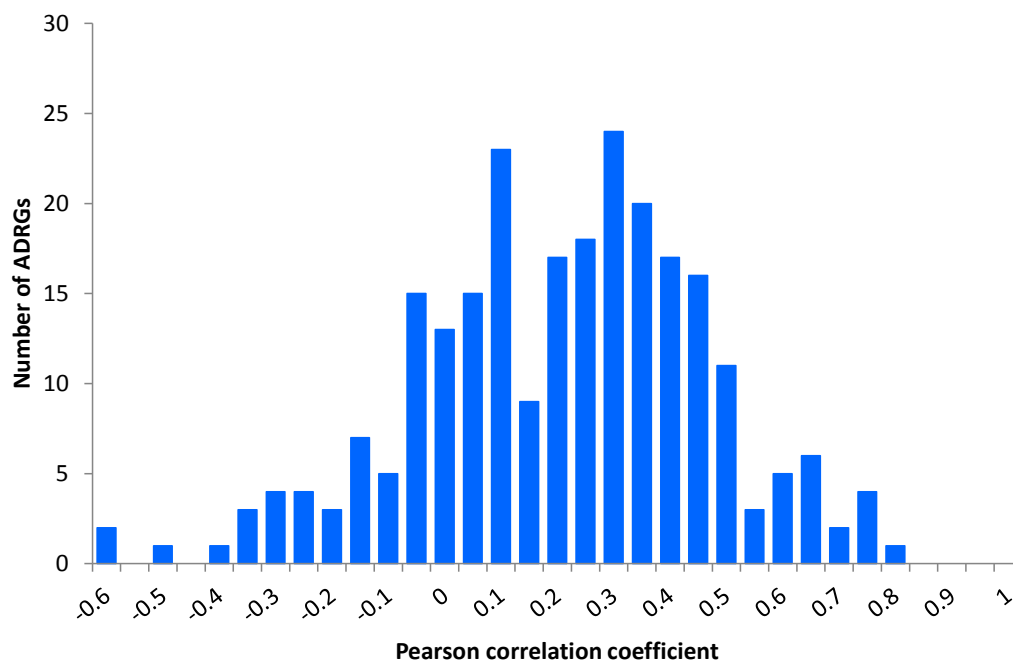
Figure 5: Scatter plot of diagnosis profile mean cost by CCL value for diseases occurring in $E_3(F14)$.



The performance of CCLs within an ADRG can be measured by the level of correlation between diagnosis profile mean costs and corresponding CCLs, and in this respect, Figure 4 and Figure 5 demonstrate that CCL assignments within $F14$ are performing poorly. Specifically, the plotted values of diagnoses with CCL values of 2, 3 or 4 are not exhibiting any distinguishing cost characteristics compared to the plotted values of diagnoses with a CCL value of 0.

This very low correlation of CCLs and diagnosis profile mean costs is strikingly evident across almost all ADRGs. For example, Figure 6 shows the distribution of Pearson correlation coefficients across ADRGs resulting from the comparison of CCLs and diagnosis profile mean costs of episodes with precisely two diagnoses (i.e. within each $E_2(A)$). These coefficients ranged between -0.6 and 0.8, with a mean value of 0.21 and a standard deviation of 0.26. Note that the correlation coefficients represented in Figure 6 have been calculated on only those diagnosis profiles containing at least 30 episodes, and furthermore, Figure 6 only includes those ADRGs with at least 100 episodes and 10 diagnosis profiles.

Figure 6: Distribution of Pearson correlation coefficients across ADRGs resulting from comparison of CCLs and diagnosis profile mean costs of episodes with two diagnoses.



The same lack of correlation is exhibited when comparisons are made among episodes with three, four or five diagnoses.

Key Finding 3

The current method of measuring case complexity, the CCLs, exhibits very little (if any) correlation with cost. This was based on an in depth review of the existing case complexity system using three years of patient level cost and activity data from 2009-10 to 2011-12.

Recommendation 1

Based on Key Findings 1 – 3, ACCD in consultation with the DTG and CCAG recommends that a new conceptually based, formally derived and data driven case complexity system be developed for AR-DRG Version 8.0 and future versions of the AR-DRG classification.



4 The Episode Clinical Complexity Model

The observed poor performance of CCLs, together with their lack of recorded conceptual foundation, prompted the development of a conceptually based, formally derived and empirically driven approach to quantifying diagnosis complexity. This in turn led to a revised approach to measuring episode clinical complexity. This section details the development of this Episode Clinical Complexity (ECC) Model.

4.1 Summary of the development of the Episode Clinical Complexity Model

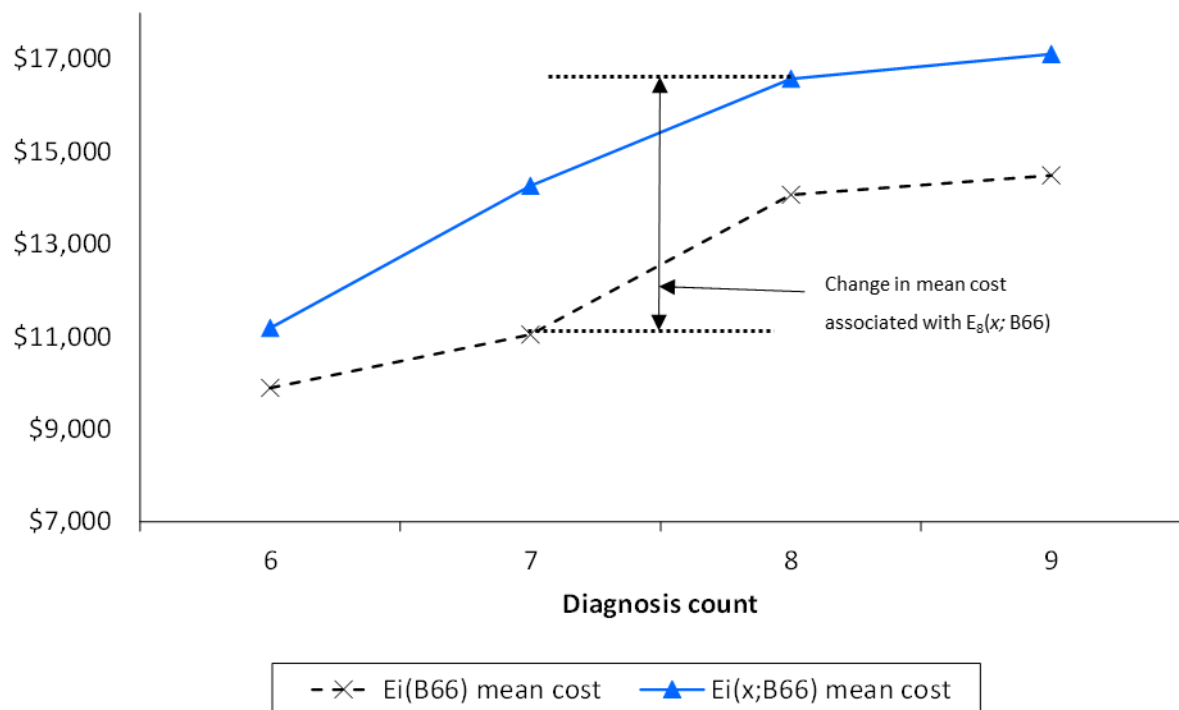
The formal development of the ECC Model is introduced by first outlining the conceptual structure of the model.

4.1.1 Measuring relative changes in cost associated with diagnosis cost profiles

Figure 3 on page 33 provided an illustration of how the cost within an ARDG varies for a particular diagnosis as the number of total diagnoses in an episode increases. The relative impact of each diagnosis may be formalised as follows. Each of the diagnosis profiles within a particular $E_i(A)$ can be considered relative to the previous ADRG profile $E_{i-1}(A)$. This provides a way of quantifying the relative cost associated with a particular diagnosis profile $E_i(x; A)$, and these relative costs can then be combined across the i taking into account the sample size of each profile.

Figure 7 illustrates how the cost profiles of $E_8(x; A)$ may be compared to the cost profile of $E_7(A)$ to derive the relative contribution of x in explaining the change in cost from $E_7(A)$ to $E_8(A)$. Essentially, the total change in mean costs associated with $E_8(x; A)$, from a diagnosis count of seven, is equal to the difference between the mean cost of $E_7(A)$ and the mean cost of $E_8(x; A)$.

Figure 7: Illustration of the change in cost associated with x in $E_8(B66)$.



Three points should be highlighted:

- The method used to quantify relative contribution to costs does not attempt to isolate the costs of individual diagnoses, nor does it attempt to establish causality between diagnosis and cost;
- Estimates of change are affected by sample size and variation of the ADRG profile of $E_7(A)$ as well as that of the diagnosis profiles within $E_8(A)$; and
- The change in cost may be measured in absolute terms or relative terms (i.e. in terms of an absolute change or a percentage change from the ADRG profile of $E_7(A)$ to the diagnosis profiles of $E_8(A)$).

To add further clarity to the first point above, the method used to measure cost associated with $E_i(x;A)$ does not distinguish between costs that may be caused by x in isolation and other costs which may be present due to correlations among appearance of x with other



diagnoses. To make clear that there is no causality being established, these costs are referred to as cost associations. The ‘overlapping’ of cost associations among diagnoses is accounted for through the process of combining the cost associations at the episode level.

The second point is addressed by deriving and using a modelled form of the $E_i(A)$ profiles based on the empirical distribution across the $E_i(A)$ instead of using the empirical distributions themselves. These models provide a best-fit trend across the $E_i(A)$ that all adhere to a specified model form.

With regard to the third point, both absolute and relative approaches were tested against the data, and both were shown to provide acceptable measurements of cost associations. The relative change approach was selected as the preferred option for the ECC Model as it provided a narrower distribution of diagnosis relative cost associations. Specifically, the absolute approach resulted in the majority of absolute changes in cost association between \$500 and \$5,000, whereas the relative change approach results in the majority of relative changes in cost association between 1.05 and 1.5 (i.e. between 5 per cent and 50 per cent increase).

The selection of a relative approach to measuring changes in cost also has implications in terms of the statistical model used to estimate change across the ADRG profiles $E_i(A)$ and the estimation method on which the relative changes are based. Specifically, the ADRG profile model is selected as a multiplicative model as opposed to an additive model, and the estimation of that model and use of it to measure changes associated with diagnosis cost profiles use geometric methods as opposed to arithmetic methods. For example, geometric means are used to estimate the ‘average’ relative change rather than the usual arithmetic mean, which would relate to an ‘average’ additive change. These issues are discussed in further detail in the *Modelling of ADRG costs* section (see page 41).



4.1.2 Standardising the diagnosis relative costs within each ADRG

Once the costs associated with each diagnosis profile $E_i(x; A)$ are combined across the ADRG, the resulting value is expressed in terms of an average, or standard, ADx in A . Consequently, costs associated with diagnoses in A are measured in terms of standard ADx. For example, a diagnosis x with a standardised cost of 3 in an ADRG A has been estimated to have a cost association three times that of an average ADx in A .

The estimation of a standard ADx within each ADRG is based on the best-fit model that is used to estimate trends across the ADRG profiles $E_i(A)$.

4.1.3 Combined diagnosis relative cost associations at the episode level

The standardised relative costs of diagnoses within ADRGs are combined across each episode in a way that can be used to estimate the overall costs of the episode. This process takes into account the correlation among diagnoses and overlapping of cost associations by diminishing the contribution of each diagnosis to the overall episode cost estimate. For example, two diagnoses each with standardised values of 1 will combine to form an episode score less than 2, and three diagnoses, each with a standardised value of 2 will combine to form an episode score less than 6.

The diminishing contribution of multiple diagnoses is empirically based. The process by which this is derived is detailed in the *Combining Diagnosis Complexity Levels across episodes and derivation of the Episode Clinical Complexity Score* section (page 62).

4.2 Formal development of the Episode Clinical Complexity Model

The formal development of the ECC model is set out in five stages:

1. Diagnosis exclusions;
2. Modelling of ADRG costs;
3. Estimation of relative costs associated with diagnoses within the context of ADRGs;
4. Derivation of the DCL; and
5. Combining DCLs across episodes and derivation of the ECCS.



4.2.1 Diagnosis exclusions

The first stage in the process of defining the ECC Model defines the scope of the model in terms of diagnoses considered relevant for DRG classification purposes. This process is guided by principles which aim to characterise the scope of the ECC model and have the effect of identifying diagnoses permitted to be assigned nonzero DCLs. The diagnoses not identified as in-scope are called exclusions, some of which are excluded unconditionally and others excluded conditionally (i.e. some diagnoses are excluded in circumstances where another diagnosis is present in the same episode).

Out of scope diagnoses are removed from the data (i.e. the diagnosis array) prior to its use in the development of the ECC Model. The ECC model scope principles are discussed in the *Guiding principles for Diagnosis Complexity Level assignment* section (page 78) and the exclusions can be found at Appendices 2 - 7.

4.2.2 Modelling of ADRG costs

The objective at the diagnosis level is to measure the change in cost associated with a particular diagnosis profile relative to the previous ADRG profile; that is, for each ADRG A , diagnosis x and positive integer i , the cost profile of $E_i(x; A)$ is compared in relative terms to the cost profile of $E_{i-1}(A)$.

Since these comparisons are relative rather than absolute, the average relative change in cost from $E_{i-1}(A)$ to $E_i(x; A)$ is equal to the geometric mean cost of $E_i(x; A)$ divided by the geometric mean cost of $E_{i-1}(A)$. However, given that the sample size of $E_{i-1}(A)$ may be small or even zero regardless of the sample size of $E_i(x; A)$, modelling is used to estimate the geometric mean cost of the $E_{i-1}(A)$ across each ADRG A .

To this end, the second stage in the process of defining the ECC Model involves the derivation of models that estimate geometric mean costs across ADRG profiles $E_i(A)$. As discussed in the *Measuring relative changes in cost associated with diagnosis cost profiles* section (page 37), a multiplicative model form is used to estimate ADRG costs by diagnosis



count. Further to this, to align with the diminished combined effect of multiple diagnoses introduced in the *Combined diagnosis relative cost associations at the episode level* section (page 40), a decay factor is introduced into the model form.

Each ADRG geometric mean cost model takes the defined form:¹⁰

$$C_i(A) := a \times b \times b^r \times b^{r^2} \times \dots \times b^{r^{i-1}}$$

$$= a \times b^{\frac{1-r^i}{1-r}}$$

Where the base cost a , change parameter b and decay parameter r are specific to each ADRG, and $a > 0$, $b \geq 1$ and $0 < r < 1$.

To illustrate with an example, if $a = \$1,000$, $b = 1.3$ and $r = 0.85$, then

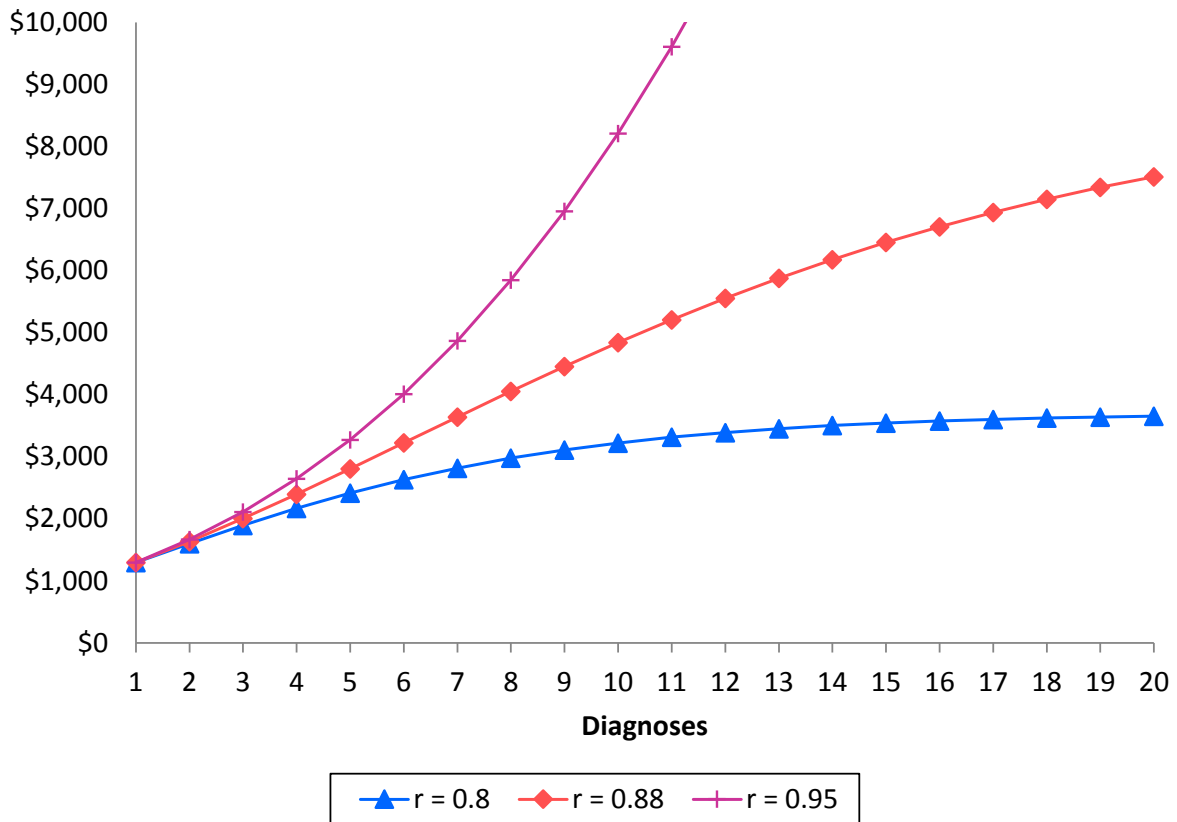
- $C_1(A) = \$1,000 \times 1.3$
- $C_2(A) = \$1,000 \times 1.3 \times (1.3)^{0.85} \approx \$1,000 \times 1.3 \times 1.25$
- $C_3(A) = \$1,000 \times 1.3 \times (1.3)^{0.85} \times (1.3)^{(0.85)^2} \approx \$1,000 \times 1.3 \times 1.25 \times 1.21$
- $C_4(A) = \$1,000 \times 1.3 \times (1.3)^{0.85} \times (1.3)^{(0.85)^2} \times (1.3)^{(0.85)^3} \approx \$1,000 \times 1.3 \times 1.25 \times 1.21 \times 1.17$

So, the first (principal) diagnosis increases the base cost of \$1,000 by 30%, the second diagnosis then increases the cost by 25%, the third by 21%, the fourth by 17%, and so on with ever decreasing levels of growth.

The parameters a , b and r together determine the model form by specifying the base level of cost, the underlying relative change in cost associated with ADx, and the diminished effect of ADx in combination. Figure 8 illustrates the way in which the decay parameter r offsets the change parameter b .

¹⁰ Throughout this report the modified equal sign “:=” is used to indicate that a definition is being made; that is, the modified equal sign “:=” can be read as “is, by definition, equal to”.

Figure 8: Illustration of $C_i(A)$ model form with $a = \$1,000$, $b = 1.3$ and three choices of r : $r = 0.8$; $r = 0.88$; and $r = 0.95$.



The geometric mean of a set of numbers can be calculated by taking the logarithm of each number in the set, taking the arithmetic mean of the resulting log values, and transforming the arithmetic mean back using exponentiation. That is, for a set of positive values $\{p_1, p_2, \dots, p_n\}$, the following equation shows two equivalent ways of calculating the geometric mean of the set:

$$\left(\prod_{i=1}^n p_i \right)^{\frac{1}{n}} = e^{\frac{\sum_{i=1}^n \ln(p_i)}{n}}.$$

Similar to this relationship between geometric and (transformed) arithmetic means, a model estimating geometric mean costs across the $E_i(A)$ of each ADRG A can be derived by taking the logarithm of all episode costs, finding a least squares best fit model of the log values, and taking the exponent of the model.



To minimise the influence of high leverage observations on the estimation of $C_i(A)$ model parameters, data on which they are derived is restricted to episodes containing less than or equal to 20 diagnosis codes; that is, $E_i(A)$ where $i \leq 20$.

Since the $C_i(A)$ models are defined to have a decayed multiplicative form

$C_i(A) = a \times b^{\frac{1-r^i}{1-r}}$, the log-transformed models have the form:

$$\begin{aligned}\ln(C_i(A)) &= \ln\left(a \times b^{\frac{1-r^i}{1-r}}\right) \\ &= \ln(a) + \frac{1-r^i}{1-r} \times \ln(b)\end{aligned}$$

Since these log-transformed models have a decayed linear (or additive) form as opposed to a simple linear form, numerical methods are used to find a least-squares best fit solution for each ADRG. The solutions are then transformed back to their original multiplicative form using exponentiation. To illustrate the result of this process, Figure 9 shows the decayed multiplicative model $C_i(G03)$ along with the geometric mean cost of each $E_i(G03)$.

Figure 9: The decayed multiplicative model $C_i(G03)$ together with the geometric mean cost of each ADRG profile $E_i(G03)$.

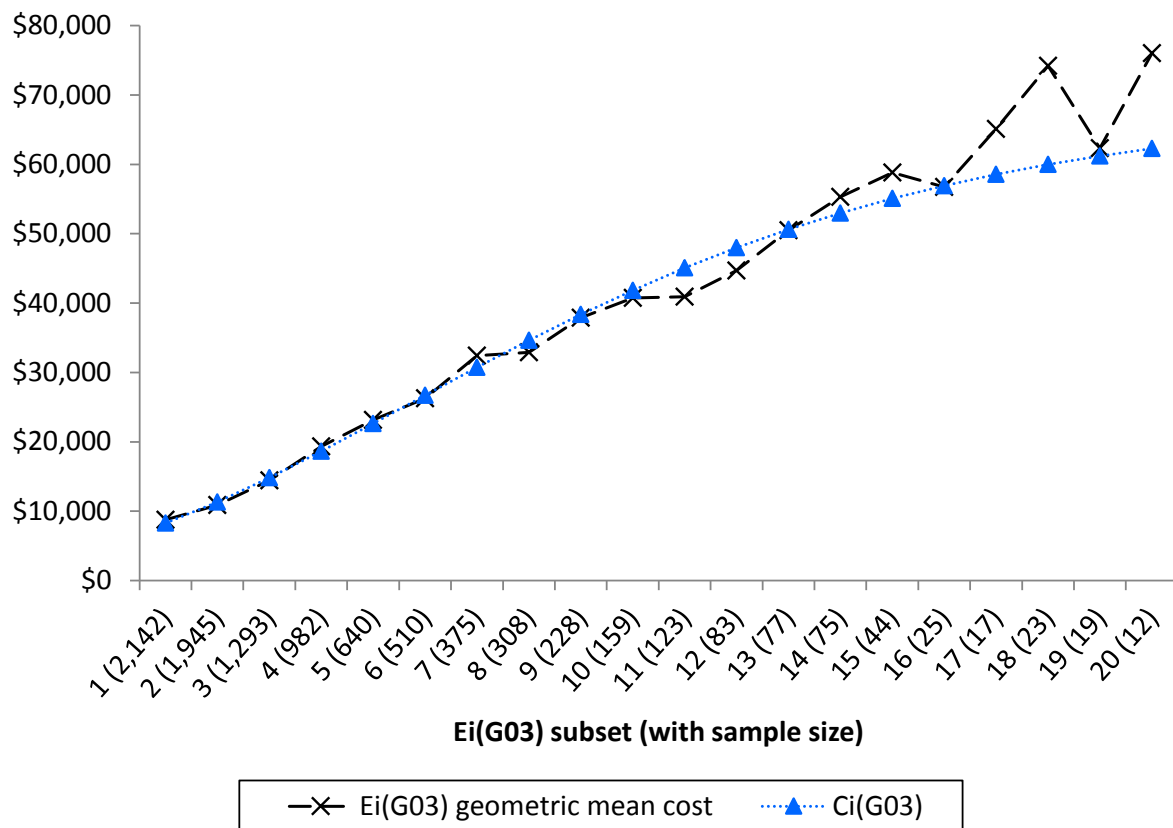


Table 2 provides the values of $C_i(G03)$, $E_i(G03)$, and the proportional variation of $C_i(G03)$ to $E_i(G03)$ for each value of i . The approximated parameters of $C_i(G03)$ are $a = \$5,710$, $b = 1.45$ and $r = 0.85$; that is,

$$C_i(G03) = \$5,710 \times (1.45)^{\frac{1-(0.85)^i}{1-0.85}}.$$

Figure 10 and Figure 11 respectively show the distributions of $C_i(A)$ relative change and decay parameters across ADRGs, and Figure 12 displays a scatter plot showing the joint distribution of the two parameters. Figure 12 shows a strong negative correlation between the relative change and decay parameters of $C_i(A)$, that demonstrates the way in which higher levels of growth are offset by higher levels of decay.



Table 2: The decayed multiplicative model $C_i(G03)$ together with the geometric mean cost of each ADRG profile $E_i(G03)$.

$E_i(G03)$ subset (sample size)	$C_i(G03)$	$E_i(G03)$ geometric mean cost	Proportional variation of $C_i(G03)$ to $E_i(G03)$
1 (2,142)	\$8,268	\$8,755	+6%
2 (1,945)	\$11,333	\$10,889	-4%
3 (1,293)	\$14,822	\$14,422	-3%
4 (982)	\$18,626	\$19,346	+4%
5 (640)	\$22,623	\$23,156	+2%
6 (510)	\$26,694	\$26,267	-2%
7 (375)	\$30,731	\$32,412	+5%
8 (308)	\$34,645	\$32,918	-5%
9 (228)	\$38,367	\$37,905	-1%
10 (159)	\$41,848	\$40,750	-3%
11 (123)	\$45,059	\$40,907	-10%
12 (83)	\$47,985	\$44,709	-7%
13 (77)	\$50,625	\$50,485	0%
14 (75)	\$52,986	\$55,307	+4%
15 (44)	\$55,082	\$58,800	+6%
16 (25)	\$56,931	\$56,746	0%
17 (17)	\$58,554	\$65,109	+10%
18 (23)	\$59,971	\$74,189	+19%
19 (19)	\$61,205	\$62,343	+2%
20 (12)	\$62,275	\$76,027	+18%

Figure 10: Distribution of $C_i(A)$ relative change parameter (i.e. b).

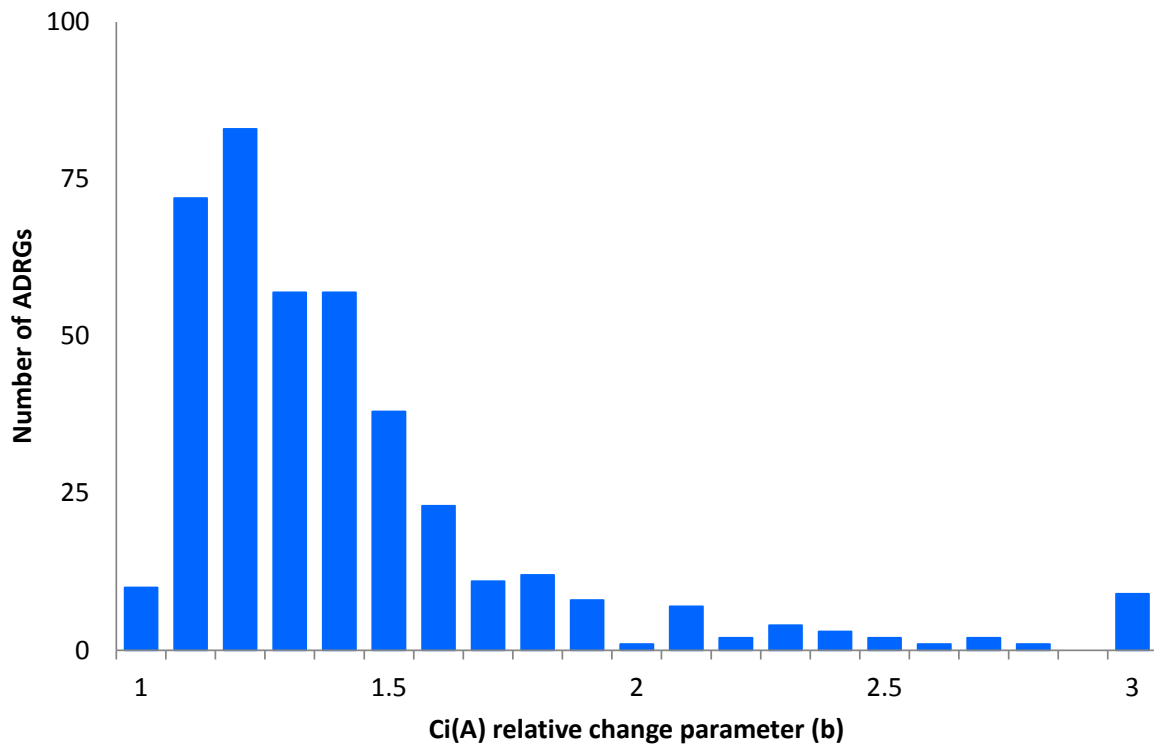


Figure 11: Distribution of $C_i(A)$ decay parameter (i.e. r).

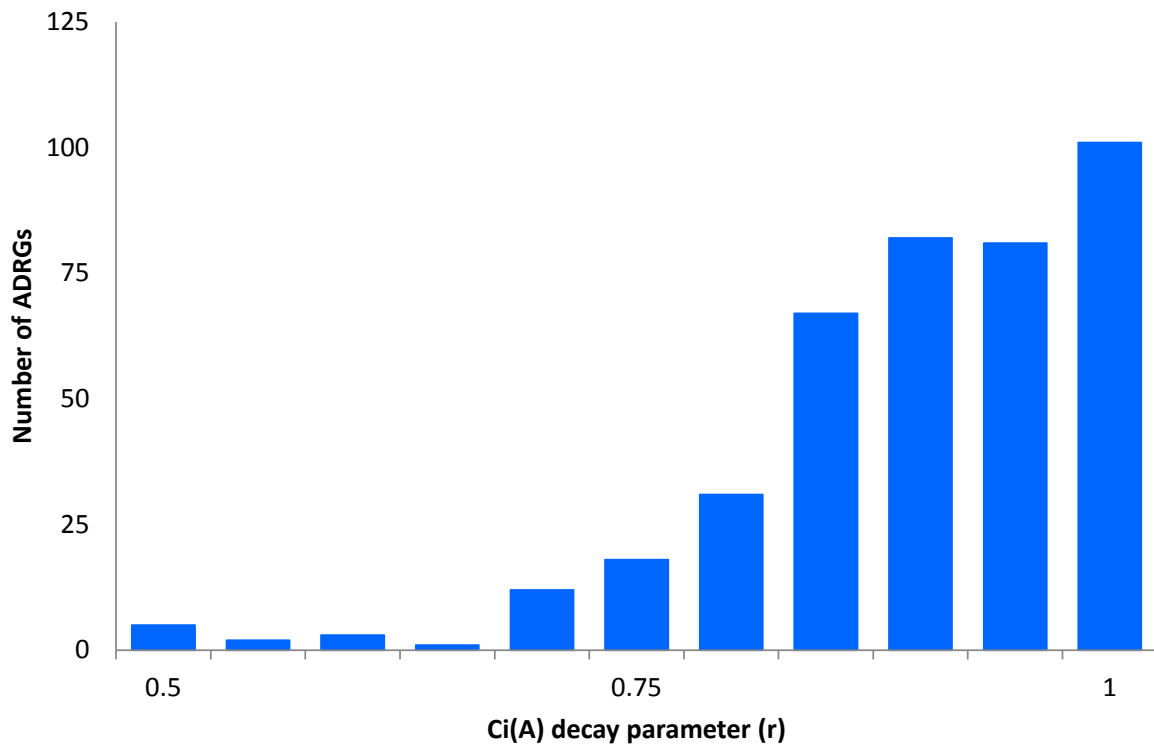
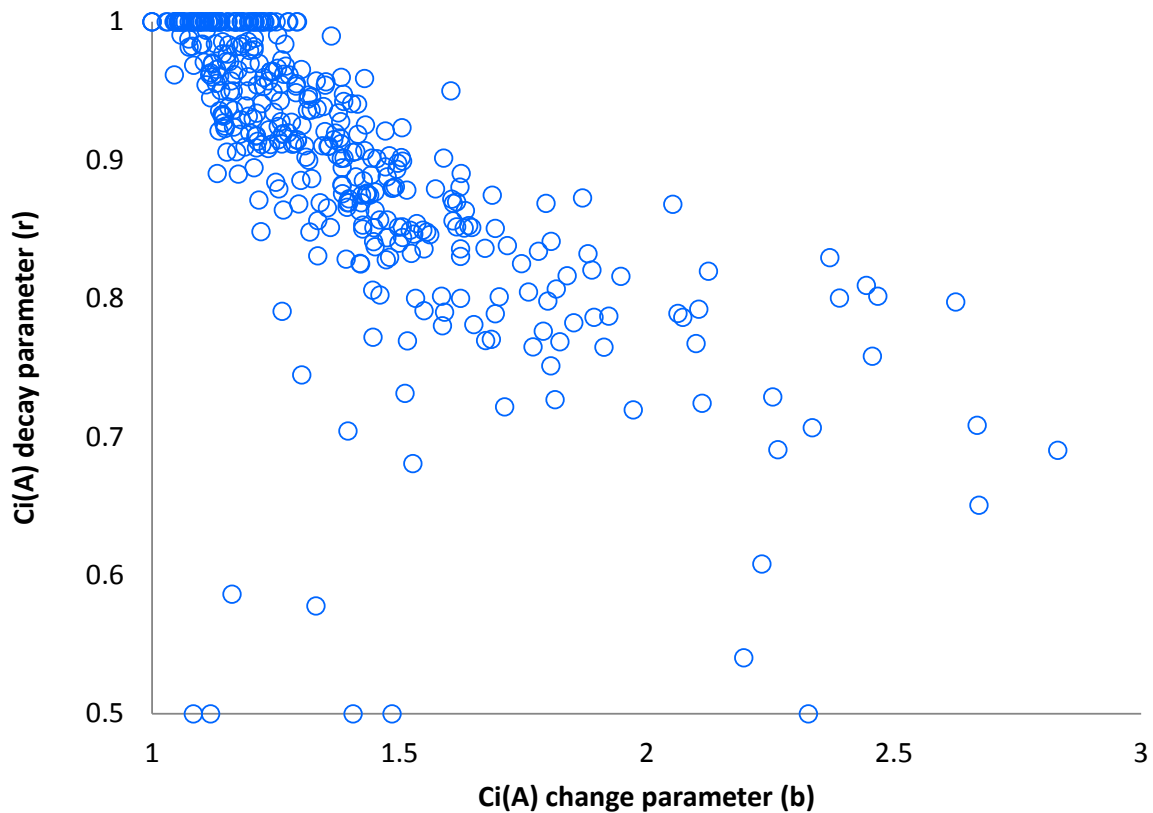


Figure 12: Scatter showing joint distribution of $C_i(A)$ relative change and decay parameters (i.e. b and r).



4.2.3 Estimation of relative costs associated with diagnoses within the context of ADRGs

The third stage in the process of defining the ECC Model is the estimation of relative costs associated with diagnoses within the context of ADRGs; that is, for a given diagnosis x and ADRG A , the estimation of the relative cost associated with x within A . This begins with the estimation of the relative cost associated with each diagnosis profile $E_i(x; A)$. These estimates are then combined across all ADRG profiles of A .

The following notation is used to express these relative cost associations formally:

1. the cost of an episode e is denoted by $c(e)$; and
2. the number of episodes in a set $E_i(x; A)$ is denoted by $|E_i(x; A)|$.



Then, for a given diagnosis x , an ADRG A and a positive integer i , the relative contribution of x in explaining the change in cost associated with $E_i(x; A)$ is defined as:

$$C_i(x; A) := \frac{\left(\prod_{e \in E_i(x; A)} c(e) \right)^{1/|E_i(x; A)|}}{C_{i-1}(A)}.$$

In other words, the relative cost associated with x in $E_i(A)$ is equal to the geometric mean cost of episodes in $E_i(x; A)$ divided by the model value $C_{i-1}(A)$. Note that $C_{i-1}(A)$ can be thought of as the modelled estimate of the geometric mean cost of episodes in $E_{i-1}(A)$.

However, $C_0(A)$, which is equal to the a parameter from $ab^{\frac{1-r^i}{1-r}}$, is used to define $C_1(x; A)$ even though $E_0(A)$ is not defined.¹¹

The relative cost associated with x in A is then defined as the geometric mean of the $C_i(x; A)$ weighted by the sample size of each $E_i(x; A)$. That is,

$$C(x; A) := \left(\prod_i C_i(x; A)^{|E_i(x; A)|} \right)^{1/\sum_i |E_i(x; A)|}.$$

As with the derivation of the $C_i(A)$ models, data is restricted to $E_i(A)$ where $i \leq 20$ when calculating $C(x; A)$ to minimise the effect of overly influential observations.

4.2.4 Derivation of the Diagnosis Complexity Level

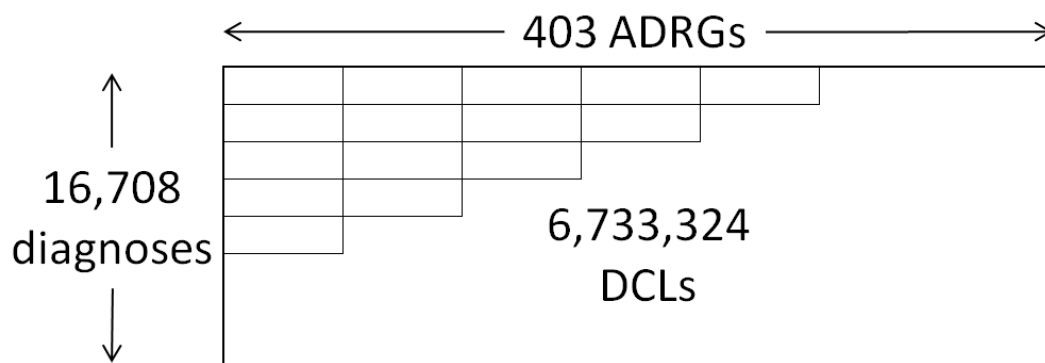
The fourth stage in the development of the ECC Model is to use the relative cost estimates $C(x; A)$ to derive the DCL of each diagnosis within each ADRG. Every possible combination of diagnosis x and ADRG A is assigned a DCL, denoted $DCL(x; A)$, based on a

¹¹ Although $E_0(A)$ has not been defined, there are episodes that could be considered elements of such a set; namely, episodes which have had all diagnoses removed as unconditional exclusions.



standardised relative cost associated with x in A . There are 16,708 diagnoses and 403 (non-error) ADRGs, which combine to give 6.7 million DCLs, illustrated by Figure 13.

Figure 13: Illustration of the DCL array.



All out of scope diagnoses (i.e. unconditionally excluded diagnoses) are assigned a DCL of zero; that is, each unconditionally excluded diagnosis x has $DCL(x; A) = 0$ for all ADRGs A . There are 4,285 out of scope diagnosis codes (see the *Guiding principles for Diagnosis Complexity Level assignment* on page 78), leaving 12,423 in-scope diagnosis codes. Therefore, there are 5,006,469 entries in the DCL array that may be nonzero.

The first step in the derivation of DCLs for in-scope diagnoses is the combining of relative cost estimates $C(x; A)$ to obtain adequate sample size and robustness of DCL estimation. To describe the way in which this occurs, a classification of diagnoses into what is termed *Coherent Diagnosis Classes (CDCs)* is required. Historically, CDCs have been referred to as Categorical ADRGs. However, to avoid confusion with the ADRG episode classification, this diagnosis classification is termed CDC.

The set of all CDCs aligns with the set of all ADRGs from the Medical Partition in the sense that diagnoses are classified into the same CDC if they appear as PDxs within the same medical ADRG. That is, for each Medical ADRG A , the set of PDxs of A fall into the same CDC. There are several in-scope diagnoses that do not appear as a PDx for any Medical ADRG, and these diagnoses are assigned to a clinically appropriate CDC. There are also 67 diagnosis codes that have sex-specific CDC assignment. Appendix 8 lists all in-scope diagnoses together with their CDC.



The level of precision of DCLs needs to be balanced against sample variation, stability over time, and performance of the resulting ECC Model. With these factors in mind, the finest level of precision of the DCLs is generally taken to be the 3-character code category within CDC. That is, all codes that belong to the same 3-character code category and belong to the same CDC are assigned the same DCL. However, a DCL is only assigned at this level if there is adequate sample size to do so. A sample size of 100 episodes is taken as the minimum threshold used for the calculation of DCLs, and relative cost estimates are combined in a specific way until this threshold is reached.

Although the 3-character code category within CDC is generally taken as the default level of DCL precision, the *Enhancing DCL precision* section (page 113) describes how this level of precision may be enhanced to the fourth or fifth character level, with the fifth character level being the highest level of precision at full code specification.

The derivation of DCLs begins with the relative change in cost estimates $C(x; A)$ defined in the previous section. However, as these estimates are combined to the level of 3-character code category within CDC, those diagnoses with sex-specific CDC assignment are split into male and female versions; that is, a diagnosis x with sex-specific CDC assignment is split into a male version x_M and a female version x_F throughout the data. This results in two sets of relative cost estimates of the form $C(x_M; A)$ and $C(x_F; A)$, which are then combined into their associated CDC groupings at the end of the DCL derivation process.

As indicated above, the process of DCL calculation is dependent on a sample size threshold of 100 episodes, and this threshold is obtained by combining relative cost estimates through an increasing hierarchy of episode groupings. There are two dimensions in which the relative cost estimates $C(x; A)$ may be aggregated: the diagnosis dimension and the ADRG dimension, with the diagnosis dimension taken as the preferred direction of aggregation.



There are seven levels of aggregation used in the diagnosis dimension (with level 3 being the default starting level):

1. diagnosis code;
2. 4-character category within CDC;
3. 3-character category within CDC;
4. Code Block within CDC;
5. Code Section within CDC;
6. Code Chapter within CDC; and
7. CDC.

There are four levels of aggregation used in the ADRG dimension:

1. ADRG;
2. Major Diagnostic Category (MDC) by Partition;
3. MDC; and
4. All ADRGs.

Figure 14 shows the way in which relative cost estimates are combined through these levels of aggregation.

Figure 14: Illustration of aggregation hierarchy for DCL calculation.

Two added levels of precision where identified as appropriate

	ADRG	MDC by Partition	MDC	All ADRGs
diagnosis code	1 ¹			
4-character category within CDC	2 ²			
3-character category within CDC	1	6		
Code Block within CDC	2	7		
Code Section within CDC	3	8		
Code Chapter within CDC	4	9	11	13
CDC	5	10	12	14



The process of combining relative cost estimates until an adequate sample is found is done in an iterative way that only brings in sufficient sample to arrive at the threshold. For example, if a particular 3-character category within CDC by ADRG combination contains 95 episodes, and the next level of aggregation (Code Block within CDC by ADRG) has 5,000 episodes (including the 95 from the previous level), then of the 4,905 extra episodes (i.e. 5,000 minus 95) only a further five episodes are used to form the combined estimate. These five extra episodes are taken to each have a relative cost estimate equal to the geometric mean of the 4,905 episodes. The way in which this calculation occurs is detailed as follows.¹²

For each diagnosis x and ADRG A , Figure 14 specifies 14 levels of aggregation that $E(x; A)$ appears within. Specifically, denoting the five levels of diagnosis aggregation containing x (starting at the level of 3-character category within CDC) by X_1, \dots, X_5 and the four levels of ADRG aggregation containing A by $A_1 (= A), \dots, A_4$, then the 14 levels of aggregation containing $E(x; A)$ are given in Table 3.

¹² Note that diagnoses with sex-specific CDC assignment are split throughout this process.

Table 3: Aggregation hierarchy for a diagnosis-ADRG combination ($x;A$).

ADRG level (A)	Diagnosis level (X)	Notation
ADRG	3-character category within CDC	$X_1 \times A_1$
	Code Block within CDC	$X_2 \times A_1$
	Code Section within CDC	$X_3 \times A_1$
	Code Chapter within CDC	$X_4 \times A_1$
	CDC	$X_5 \times A_1$
MDC by Partition	3-character category within CDC	$X_1 \times A_2$
	Code Block within CDC	$X_2 \times A_2$
	Code Section within CDC	$X_3 \times A_2$
	Code Chapter within CDC	$X_4 \times A_2$
	CDC	$X_5 \times A_2$
MDC	Code Chapter within CDC	$X_4 \times A_3$
	CDC	$X_5 \times A_3$
All ADRGs	Code Chapter within CDC	$X_4 \times A_4$
	CDC	$X_5 \times A_4$

In order to derive a combined relative cost associated with a diagnosis x in an ADRG A , the weighted geometric mean of the $C(x;A)$ at each level of aggregation specified in Table 3 is first calculated.

$$C(X_i;A_j) := \left(\prod_{y \in X_i, B \in A_j} C(y;B)^{|E(y;B)|} \right)^{1/\sum_{y \in X_i, B \in A_j} |E(y;B)|}.$$



These $C(X_i; A_j)$ are used to iteratively calculate a cumulative relative cost estimate for x in A through the 14 levels until the sample size threshold of 100 is reached. The $C(X_i; A_j)$ are denoted by C_k , where k is the corresponding level of the aggregation hierarchy. Table 4 specifies this notation.

Table 4: Notation used in calculation of cumulative relative cost.

ADRG level (A)	Diagnosis level (X)	Notation	Defined as
ADRG	3-character category within CDC	$C(X_1; A_1)$	C_1
	Code Block within CDC	$C(X_2; A_1)$	C_2
	Code Section within CDC	$C(X_3; A_1)$	C_3
	Code Chapter within CDC	$C(X_4; A_1)$	C_4
	CDC	$C(X_5; A_1)$	C_5
MDC by Partition	3-character category within CDC	$C(X_1; A_2)$	C_6
	Code Block within CDC	$C(X_2; A_2)$	C_7
	Code Section within CDC	$C(X_3; A_2)$	C_8
	Code Chapter within CDC	$C(X_4; A_2)$	C_9
	CDC	$C(X_5; A_2)$	C_{10}
MDC	Code Chapter within CDC	$C(X_4; A_3)$	C_{11}
	CDC	$C(X_5; A_3)$	C_{12}
All ADRGs	Code Chapter within CDC	$C(X_4; A_4)$	C_{13}
	CDC	$C(X_5; A_4)$	C_{14}



Although the process of deriving a combined relative cost estimate for x in A is undertaken sequentially through the 14 levels of Table 3, there is not a complete ordering of the hierarchy in terms of set containment of the $X_i \times A_j$. For example, although the $X_i \times A_j$ of levels 1 through 5 are sequentially contained in each other, the $X_i \times A_j$ of level 6 does not contain that of level 5; that is, $X_1 \times A_2$ does not contain $X_5 \times A_1$. Instead, level 5 and level 6 have only level 1 episodes in common; that is,

$$(X_5 \times A_1) \cap (X_1 \times A_2) = X_1 \times A_1.$$

In general, $X_i \times A_j$ contains $X_k \times A_l$ only when both $k \leq i$ and $l \leq j$.

To account for this partial containment of consecutive levels, the $X_i \times A_j$ of each level is restricted to those episodes not already present in previous levels of the hierarchy. This results in a sequence of episode sets $\{E_i\}_{i=1}^{14}$ that are mutually exclusive (i.e. non-intersecting) and that align with the cumulative process of defining the combined relative cost of x in A . Table 5 specifies how each of the E_i are defined.

**Table 5: Definition of the sequence of mutually exclusive E_i episode sets.**

ADRG level (A)	Diagnosis level (X)	Notation	Defined as
ADRG	3-character category within CDC	$X_1 \times A_1$	E_1
	Code Block within CDC	$(X_2 \times A_1) \setminus E_1$	E_2
	Code Section within CDC	$(X_3 \times A_1) \setminus (E_1 \cup E_2)$	E_3
	Code Chapter within CDC	$(X_4 \times A_1) \setminus (E_1 \cup E_2 \cup E_3)$	E_4
	CDC	$(X_5 \times A_1) \setminus (\cup_{i=1}^4 E_i)$	E_5
MDC by Partition	3-character category within CDC	$(X_1 \times A_2) \setminus (\cup_{i=1}^5 E_i)$	E_6
	Code Block within CDC	$(X_2 \times A_2) \setminus (\cup_{i=1}^6 E_i)$	E_7
	Code Section within CDC	$(X_3 \times A_2) \setminus (\cup_{i=1}^7 E_i)$	E_8
	Code Chapter within CDC	$(X_4 \times A_2) \setminus (\cup_{i=1}^8 E_i)$	E_9
	CDC	$(X_5 \times A_2) \setminus (\cup_{i=1}^9 E_i)$	E_{10}
MDC	Code Chapter within CDC	$(X_4 \times A_3) \setminus (\cup_{i=1}^{10} E_i)$	E_{11}
	CDC	$(X_5 \times A_3) \setminus (\cup_{i=1}^{11} E_i)$	E_{12}
All ADRGs	Code Chapter within CDC	$(X_4 \times A_4) \setminus (\cup_{i=1}^{12} E_i)$	E_{13}
	CDC	$(X_5 \times A_4) \setminus (\cup_{i=1}^{13} E_i)$	E_{14}



The cumulative relative change in cost associated with x in A , denoted $\bar{C}(x; A)$, is defined iteratively as follows:

Step 1:

- a. If $|E_1| \geq 100$, then define $\bar{C}(x; A) := C_1$. In this case, calculation of $\bar{C}(x; A)$ is complete.
- b. If $|E_1| < 100$, then define $\bar{C}_1 := C_1$. In this case, proceed to Step 2.

Step 2:

- a. If $|E_1| + |E_2| \geq 100$, then define $\bar{C}(x; A) := \frac{|E_1| \times \bar{C}_1 + (100 - |E_1|) \times C_2}{100}$. In this case, calculation of $\bar{C}(x; A)$ is complete.
- b. If $|E_1| + |E_2| < 100$, then define $\bar{C}_2 := \frac{|E_1| \times \bar{C}_1 + |E_2| \times C_2}{|E_1| + |E_2|}$. In this case, proceed to Step 3.

This process continues, with Step n defined as follows:

Step n :

- a. If $\sum_{i=1}^n |E_i| \geq 100$, then define $\bar{C}(x; A) := \frac{\left(\sum_{i=1}^{n-1} |E_i|\right) \times \bar{C}_{n-1} + \left(100 - \left(\sum_{i=1}^{n-1} |E_i|\right)\right) \times C_n}{100}$. In this case, calculation of $\bar{C}(x; A)$ is complete.
- b. If $\sum_{i=1}^n |E_i| < 100$, then define $\bar{C}_n := \frac{\left(\sum_{i=1}^{n-1} |E_i|\right) \times \bar{C}_{n-1} + |E_n| \times C_n}{\sum_{i=1}^n |E_i|}$. In this case, proceed to Step $n+1$.

If this process continues to $n = 14$ and the sample size threshold is not satisfied at that step, then define $\bar{C}(x; A) := 1$.



Each diagnosis x with sex-specific CDC assignment has its cumulative relative change estimate $\bar{C}(x; A)$ defined as the average of $\bar{C}(x_M; A)$ and $\bar{C}(x_F; A)$ weighted by the percentage split of episodes containing x among males and females, where x has a sample size of 10 or more; otherwise $\bar{C}(x; A)$ is defined as the unweighted average of $\bar{C}(x_M; A)$ and $\bar{C}(x_F; A)$.

Table 6 summarises the levels of aggregation at which the $\bar{C}(x; A)$ are calculated.

Table 6: Summary of aggregation levels at which the $\bar{C}(x; A)$ are calculated.

Level of aggregation (diagnosis dimension)	ADRG	MDC by Partition	MDC	All ADRGs
3-character category within CDC	2.84%	10.98%	Not applicable	Not applicable
Code Block within CDC	6.22%	13.93%	Not applicable	Not applicable
Code Section within CDC	4.29%	5.14%	Not applicable	Not applicable
Code Chapter within CDC	7.54%	7.53%	10.27%	11.38%
CDC	9.91%	8.35%	1.58%	0.06%

Finally, the DCL of a diagnosis x in an ADRG A is defined by standardising $\bar{C}(x; A)$ using the modelled relative change in cost associated with ADx in A ; that is, standardising $\bar{C}(x; A)$

using the b parameter from the ADRG cost profile model $C_i(A) = ab^{\frac{1-r^i}{1-r}}$.

This standardisation process expresses the relative change estimate $\bar{C}(x; A)$ in terms of the overall relative change in cost associated with ADx in A . For example, a standardised relative change of 3 would be equivalent to the relative change associated with three average, or 'standard', ADx within A .



The standardisation of one relative change value v against another w is equal to the ratio of the logarithms of the values; i.e. $\ln v / \ln w$. To see this, note that the standardised relative change of v is the number $s(v)$ satisfying $v = w^{s(v)}$. Solving this equation gives $s(v) = \ln v / \ln w$.

The standardisation factor of $\bar{C}(x; A)$ is taken as the maximum of b , 1.1 and $1 + \$250 / a$. This maximum ensures that the standardisation factor is sufficiently greater than 1 (noting that the standardisation factor must be strictly greater than 1). The resulting DCL is then rounded to the nearest integer between 0 and 5.

Specifically, denoting the base and relative change parameters of $C_i(A)$ by a_A and b_A , the DCL of x in A is defined as

$$DCL(x, A) := \min \left(5, \max \left(0, \text{round} \left(\frac{\ln(\bar{C}(x; A))}{\ln(\max(b_A, 1.1, 1 + \$250 / a_A))} \right) \right) \right)$$

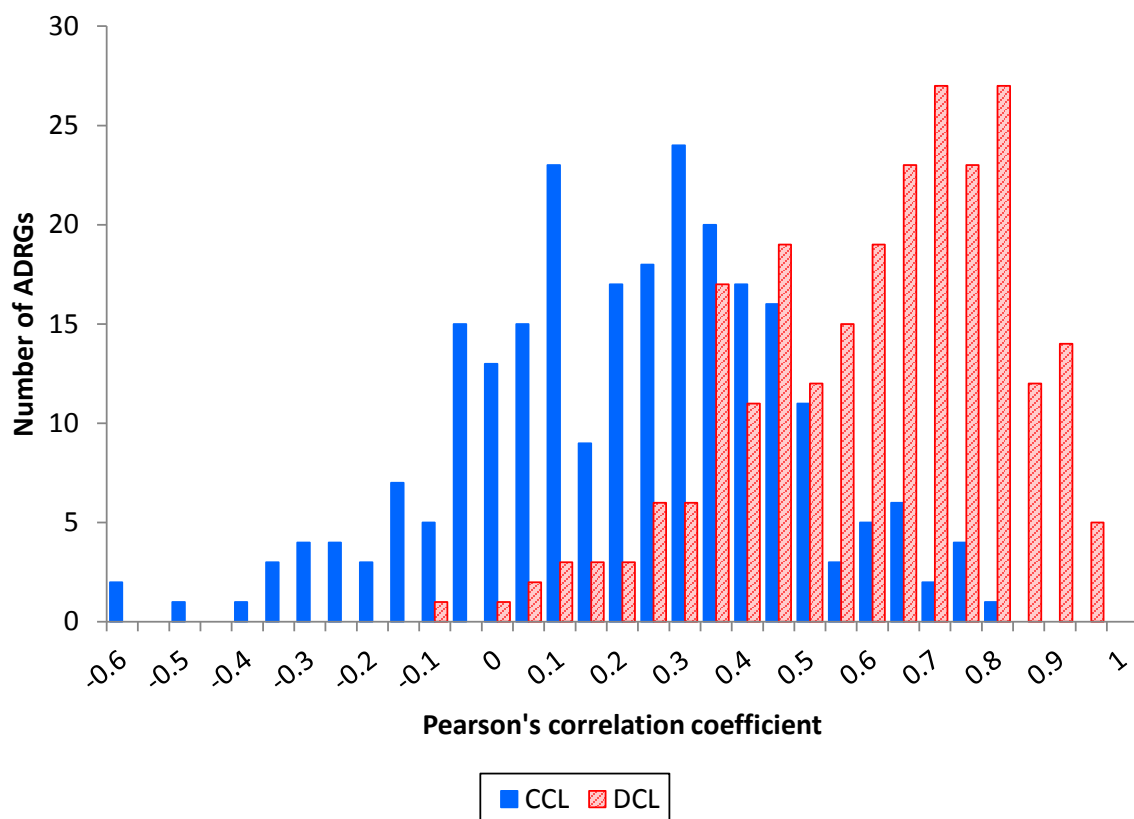
Table 7 provides a summary of DCLs across all diagnosis-ADRG combinations (excluding error ADRGs).

Table 7: Proportion of DCLs by diagnosis group.

DCL	In-scope diagnoses	Out of scope diagnoses	Total
0	22.16%	25.65%	47.81%
1	32.42%	Not applicable	32.42%
2	12.09%	Not applicable	12.09%
3	4.57%	Not applicable	4.57%
4	1.88%	Not applicable	1.88%
5	1.23%	Not applicable	1.23%
Total	74.35%	25.65%	100.00%

The DCLs exhibit a high level of correlation with episode costs within each of the ADRG profiles $E_i(A)$. Extending on the information of Figure 6, Figure 15 compares the correlation of DCLs with episode costs to the correlation of existing CCLs with episode costs among episodes with precisely two diagnoses (i.e. $E_2(A)$ across all ADRGs A). It shows DCLs to perform significantly better than CCLs in terms of correlation with costs associated with diagnoses. For example, the DCL Pearson correlation coefficients ranged between -0.10 and 0.95, with a mean value of 0.60 and a standard deviation of 0.21. This is compared to the CCL coefficients, which ranged between -0.6 and 0.8, have a mean value of 0.21 and a standard deviation of 0.26. Similar results hold for ADRG profiles with greater numbers of diagnoses (i.e. $E_3(A)$, etc.).

Figure 15: Comparison of CCL and DCL correlations with cost among episodes with exactly two diagnoses.





Key Finding 4

As a measurement of diagnosis complexity, the new conceptually based and formally defined DCLs were shown to exhibit significantly higher correlation with costs within ADRGs compared to CCLs.

Recommendation 2

Based on Key Finding 4, ACCD in consultation with the DTG and CCAG recommends that the DCL measure of diagnosis complexity be adopted as part of a new case complexity system.

4.2.5 Combining Diagnosis Complexity Levels across episodes and derivation of the Episode Clinical Complexity Score

The final stage in the development of the ECC Model is the derivation of the ECCS, which estimates the episode-level combined effect of DCLs.

For a diagnosis x and ADRG A , $DCL(x, A)$ is a standardised estimate of change in cost associated with x in A , where the standardisation is calculated with respect to the

parameter in $C_i(A) = ab^{\frac{1-r^i}{1-r}}$. From this perspective, $DCL(x, A)$ can be thought of as the index of b in $C_i(A)$. For example, the cost of an episode e in ADRG A with a single diagnosis x can be estimated by

$$c(e) \approx ab^{DCL(x, A)}.$$

The combination of multiple DCLs across a single episode can be thought of similarly, with the inclusion of a decay component that adjusts for the diminished contribution of multiple diagnoses vis-à-vis their individual contributions. For example, the cost of an episode e in ADRG A with two diagnoses x_1 and x_2 can be estimated by

$$\begin{aligned} c(e) &\approx ab^{DCL(x_1, A)} b^{r \cdot DCL(x_2, A)} \\ &= ab^{DCL(x_1, A) + r \cdot DCL(x_2, A)} \end{aligned}$$



This is generalised to n diagnoses x_1, \dots, x_n as

$$c(e) \approx ab \sum_{i=1}^n r^{i-1} \cdot DCL(x_i, A)$$

This expression was modelled across the costs of all ADRGs to find a single decay factor r that creates the least-squares best fit across all ADRGs. Each episode's diagnoses are ranked by descending DCL value for the purpose of evaluating the above expression.

A decay factor of $r = 0.84$ was identified as providing the best fit. The ECCS is then defined as the index of the resulting expression, as follows.

4.2.6 ECCS formula

The ECCS of an episode e in an ADRG A with diagnoses listed in descending order of their DCL values as x_1, \dots, x_n (i.e. $DCL(x_1, A) \geq DCL(x_2, A) \geq \dots \geq DCL(x_n, A)$) is defined as

$$ECCS(e) := \sum_{i=1}^n DCL(x_i, A) \cdot (0.84)^{i-1}.$$

Note that Conditional Exclusions (CEs) are applied to all episodes prior to evaluating this expression.

To compare and contrast with the AR-DRG V7.0 PCCL formula, Figure 16 is an extract of the AR-DRG V7.0 manual specifying the current PCCL formula.

Figure 16: Extract of AR-DRG V7.0 manual specifying PCCL formulation.

Formula for calculating a patient's clinical complexity level (PCCL)

If $\{CCL(i)\}$ is the sorted list of CCL values that remains at the end of the recursive exclusion process, the formula for calculating PCCL is:

$$PCCL = \begin{cases} 0 & \text{if there is no SDX} \\ 4 & \text{if } x > 4 \\ x & \text{otherwise} \end{cases}$$

$$\text{where } x = \text{round} \left\{ \frac{\ln \left\{ 1 + \sum_{i=k} CCL(i) \times \exp(-\alpha \times (i-k)) \right\}}{\ln(3/\alpha) / 4} \right\}$$

α is a parameter and is currently set equal to 0.4

$k = 1$ for ADRGs A06, O01-O66, P01-P68 or $k = 2$ for all other ADRGs.

Taking the majority case of $k = 2$, the PCCL formula shown in Figure 16 can be simplified as follows:

$$\begin{aligned} x &= \text{round} \left(\frac{\ln \left(1 + \sum_{i=2}^n CCL(i) \cdot e^{-0.4 \cdot (i-2)} \right)}{\ln(3/0.4) / 4} \right) \\ &\approx \text{round} \left(\frac{\ln \left(1 + \sum_{i=2}^n CCL(i) \cdot (e^{-0.4})^{i-2} \right)}{0.5} \right) \\ &\approx \text{round} \left(2 \cdot \ln \left(1 + \sum_{i=2}^n CCL(i) \cdot (0.67)^{i-2} \right) \right) \end{aligned}$$

In terms of the comparative quantification of episode complexity, this expression is similar

to an expression of the form $PCCL(e) = \sum_{i=2}^n CCL(i) \cdot (0.67)^{i-2}$. The PCCL formula and the

ECCS formula can now be seen to have two main differences:



1. The omission of a CCL for the PDx in the PCCL formula (i.e. $i = 2$ rather than $i = 1$); and
2. A PCCL decay factor of 0.67 compared to an ECCS decay factor of 0.84.

Figure 17 and Figure 18 illustrate the significantly greater correlation between ECCS and cost compared to the originally explored relationship between ADRG profiles $E_i(A)$ and cost. Of particular note is the difference in scale depicted in the y-axis of each figure, with the mean cost of Figure 17 ranging from around \$3,000 to \$13,000 and the mean cost of Figure 18 ranging from \$3,000 to \$45,000.

Figure 17: Episode costs by diagnosis count - E62 ADRG.

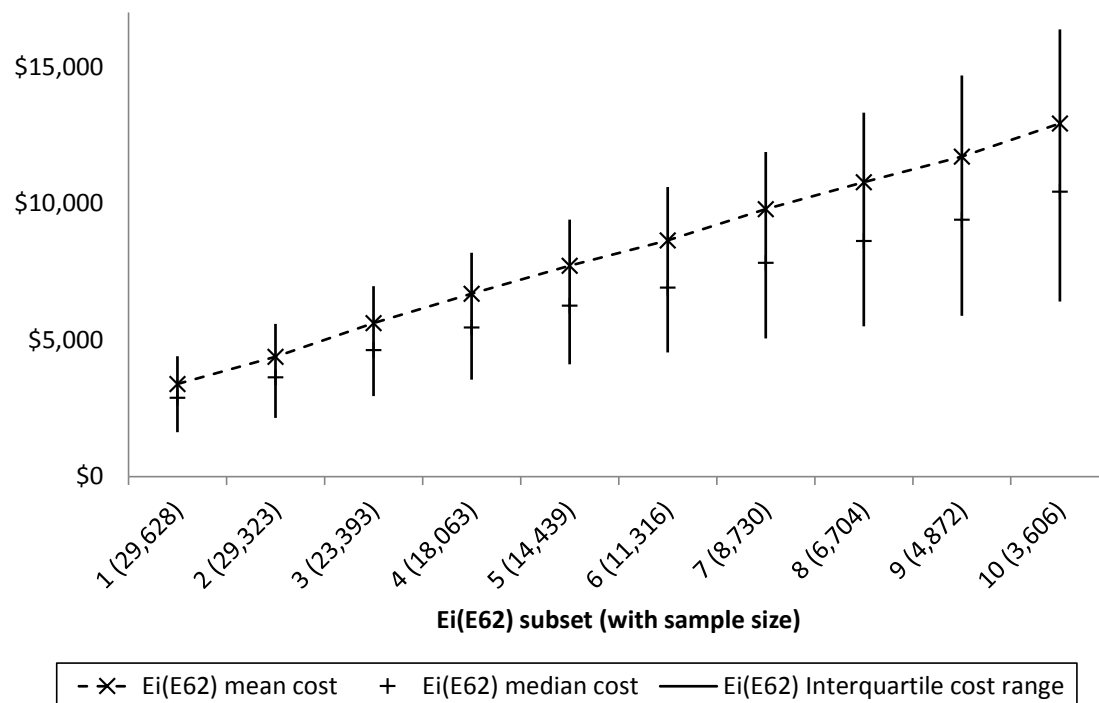
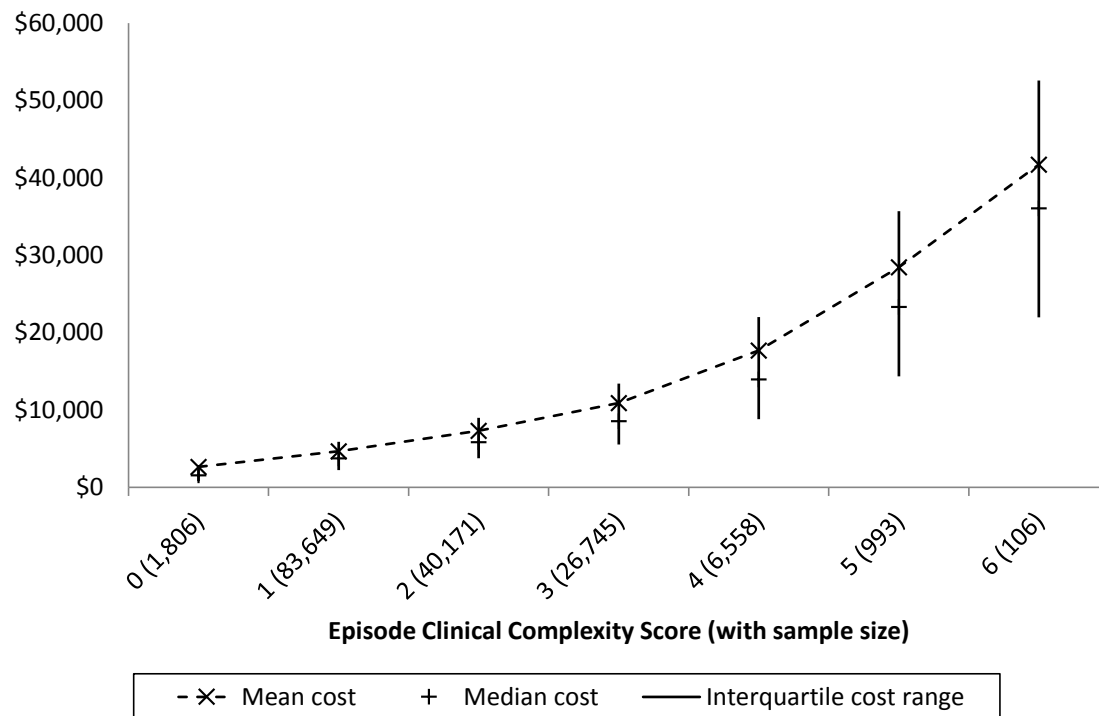


Figure 18: Episode costs by ECCS - E62 ADRG¹³



Given that the decay component of the ECCS is of the form $\sum_i r^{i-1}$, which is a geometric series, it has a natural upper bound of $\frac{1}{1-r}$. This combined with a DCL maximum of 5 and $r = 0.84$ means that the maximum ECCS possible is $5 \times \left(\frac{1}{1-0.84} \right) = 5 \times 6.25 = 31.25$. Although, across the three years of cost data from 2009-10 to 2011-12, the maximum value of the ECCS is 30.86, and less than 0.03 per cent of episodes have an ECCS of 20 or greater. Table 8 shows a percentage breakdown of episodes by ECCS, where ECCS is rounded to the nearest 0.5 for an ECCS less than 6, rounded to the nearest integer for ECCS between 6 and 10.5 and otherwise allocated to the 11+ category.

Table 8 shows that while the distribution of ECCS across episodes has a long tail of higher values, the large majority of episodes (over 97 per cent) have an ECCS of 5 or less.

¹³ Note that this sample includes ADRG profiles with eleven or more diagnoses. These have been excluded from the previous example (i.e. episode cost by diagnoses count figure).



Although the unrounded version of ECCS takes over 6,000 different values across the three years of data, the range of possible small ECCS values is limited. For example, there are only seven possible ECCS values of 3 or less; namely 0, 1, 1.84, 2, 2.5456, 2.84 and 3.

Table 8: Percentage breakdown of episodes by rounded ECCS.

Rounded ECCS	Proportion
0.0	42.97%
1.0	28.79%
2.0	12.35%
2.5	3.65%
3.0	4.38%
3.5	2.02%
4.0	1.37%
4.5	1.12%
5.0	0.72%
5.5	0.61%
6.0	0.51%
7.0	0.50%
8.0	0.33%
9.0	0.21%
10.0	0.13%
11+	0.35%
Total	100.00%



5 Treatment of the Principal diagnosis in classification design

In order to understand the role of the PDx in classification design it is essential that a description of the DRG classification process be provided in the first instance. This section describes the differences in the role of the PDx in medical, other and surgical ADRGs and provides evidence of the importance of extending its role to explaining episode clinical complexity.

5.1 The DRG Classification Process

ICD-10-AM and ACHI are the building blocks for AR-DRGs.

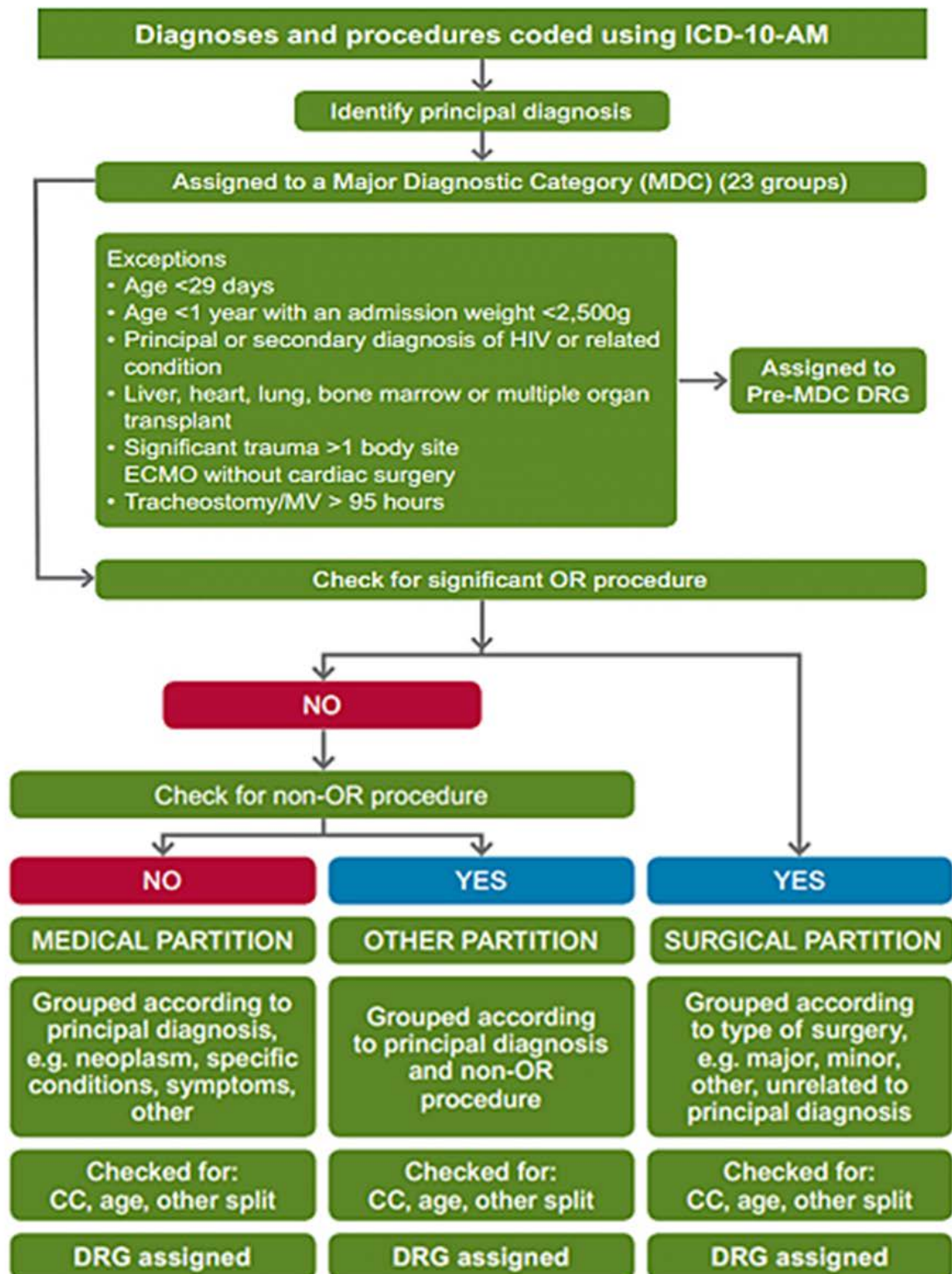
The DRG classification considers the ICD-10-AM/ACHI codes and other patient related information in the process of allocating a DRG to the specific acute admitted episode of care. A description of the current DRG classification system is included here as important background to the ACCD's case complexity review. Figure 19 on page 26 depicts the DRG classification process.

For some MDCs and ADRGs there currently are variables, other than ICD-10-AM/ACHI codes, which may affect the partitioning of the ADRGs to their component DRGs. These variables include:

- age;
- sex;
- LOS;
- same day status;
- admission weight for infants aged <365 days;
- mental health legal status; and
- mode of separation.



Figure 19: DRG classification process.



Source: *The Good Clinical Documentation Guide*. 2003



AR-DRG V7.0 incorporates 771 DRGs of which there are 768 non-error DRGs that have 403 (non-error) ADRGs, most of which are organised into 23 MDCs, generally based on body system. There are also 3 medical DRGs comprising 3 ADRGs that make up the remaining error DRGs in the 900 series. The ADRGs of each MDC are sub-divided into a maximum of three separate partitions, for surgical, other and medical. The presence or absence of 'significant' operating room (OR) and Non-OR procedures is generally responsible for the assignment of an episode of care to one of these partitions:

- surgical (336 DRGs within 188 ADRGs);
- other (48 DRGs within 27 ADRGs);
- medical (384 (non-error) DRGs within 188 (non-error) ADRGs);and
- medical error (3 DRGs within 3 ADRGs) including:
 - ungroupable;
 - unacceptable principal diagnosis; and

inconsistent neonate diagnoses. Pre-MDC processing occurs prior to allocation of the episode to a MDC. The pre-MDC process has two functions:

- Identification and assignment to one of 11 high cost DRGs that comprise the pre-MDC category.

Alters MDC assignment, where the MDC is not defined exclusively on the basis of PDx. The following information is established during pre-MDC processing:

- Was there a transplant (e.g. liver, lung, heart, bone marrow or kidney)?
- Was extracorporeal membrane oxygenation (ECMO), without cardiac surgery performed?
- Was there treatment of significant trauma at more than one body site?
- Was the patient <28 days old, or aged <1 year with admission weight <2500 grams?
- Was there a PDx related to Human Immunodeficiency Virus (HIV) with an ADx of HIV?
- Was a tracheostomy performed or did mechanical ventilation occur for >95 hours?
- Did the episode involve a ventricular assist device, spinal infusion device or neurostimulator device?



5.2 Did the diagnoses include acute quadriplegia or paraplegia? The role of the Principal diagnosis in AR-DRG development

One of the overarching principles in the design of the AR-DRG classification is that its development needs to be undertaken with maximum regard to the clinical characteristics of the patient, and with minimum regard to who is providing the service or the setting in which it is provided. However, the driver in surgical DRG partitioning is not the PDx, nor is the PDx generally considered part of the PCCL model.

With reference to Figure 19 above, there are differences in the role of the PDx in medical, other and surgical ADRGs. One of the first steps in the grouping process following MDC processing (including pre-MDC processing) is the determination of whether an episode of care has had a significant OR procedure. The following definitions of OR and NonOR procedures only relate to their use within AR-DRGs and should not be considered true clinical definitions:

- **Operating room (OR) procedures** are considered significant for all MDCs in AR-DRG V7.0. If an OR procedure is not related to the MDC that the episode is assigned to, it will be grouped to one of the unrelated OR DRGs 801A, 801B and 801C.
- **Non-operating room (NonOR) procedures** are considered significant by the classification for only some MDCs. If a NonOR procedure is not significant in an MDC, and there is no OR procedure present, the episode will be grouped to a medical DRG. Procedures identified as NonOR for purposes of AR-DRG V7.0 may actually take place in operating rooms.

Acute admitted episodes falling into surgical ADRGs are grouped according to the type of surgery, (e.g. major, minor, other or unrelated to the PDx). Given that surgery is the driver of ADRG assignment, the PDx information is not directly used in classifying surgical ADRGs. However, it is important to note that all diagnoses, no matter where in the sequence (PDx or ADx) do have an impact on resource consumption within each ADRG as suggested by empirical data. For surgical DRGs, including the PDx information in the complexity calculation allows true cost differences associated with different PDx to be captured in the DRG design.



One could argue that the PDx, specifically in medical ADRGs already has a designated role in driving the episode into a particular ADRG. However the PDx in the current AR-DRG classification's CC system allocates the episode to an ADRG but generally does not contribute to the measurement of complexity; so all PDxs within an ADRG are in effect treated as equally complex.

In the current system, many DRGs have multiple principal diagnoses, often with different associated costs but with insufficient volumes to warrant splitting the DRG. Including the PDx in complexity calculations overcomes this issue. Moreover, the current system does not recognise all disease information contained in many principal diagnoses codes. For example, the underlying classification (ICD-10-AM) contains many pre-coordinated codes (i.e. many codes contain more than one clinical concept). This means that one or more of these concepts (e.g. underlying cause, manifestation(s)) in a pre-coordinated PDx could be a contributor to cost but not considered in episode clinical complexity. See Table 9 below for examples.

In short, the current CC system discards PDx information once the ADRG is determined. This results in lost information which can have an impact on cost variation within an ADRG. The next section quantifies this variation.



Table 9: Examples of patient diagnoses within pre-coordinated codes with multiple clinical concepts.

Pre-coordinated code	Patient diagnosis	Number of clinical concepts
E50.3	Vitamin A deficiency with corneal ulceration and xerosis	3
G06.0	Intracranial abscess and granuloma	2
I11.0	Hypertensive heart disease with (congestive) heart failure	3
I72.5	Aneurysm and dissection of other precerebral arteries	2
I83.2	Varicose veins of lower extremities with both ulcer and inflammation	3
K25.6	Gastric ulcer, chronic or unspecified with both haemorrhage and perforation	3
K57.01	Diverticulosis of small intestine with haemorrhage, perforation and abscess	4
K71.7	Toxic liver disease with fibrosis and cirrhosis of liver	3
M80.45	Drug-induced osteoporosis with pathological fracture, pelvic region and thigh	3
N70.9	Salpingitis and oophoritis, unspecified	2



5.3 Principal diagnosis impact on cost

The review of the current CCLs has been undertaken by profiling the costs (and LOS) of ADRGs and diagnoses by using a method that allows comparative assessments to be made using associations between diagnoses and costs (or LOS) within each ADRG.

In the proposed ECC Model, a DCL value is assigned to each diagnosis occurring within an ADRG, regardless of whether it is a principal or additional (secondary) diagnosis. The DCL estimates the level to which the diagnosis is associated with costs, compared to that of a standard, or average, ADx. For example, a diagnosis with a DCL of 3 within an ADRG is associated with costs that are three times higher than that of an average ADx within the ADRG. This effect was observable in a large number of cases as illustrated in the following figures.

The figures below provide examples of principal diagnoses demonstrating significantly higher cost associations across their ADRGs. They show that the PDx clearly plays a role in explaining cost variation within the ADRG level of the AR-DRG classification.

Figure 20 and Figure 21 relate to two surgical ADRGs, specified by a particular set of significant OR procedures.

Figure 22 and Figure 23 relate to two medical ADRGs, the first (G67) specified by a particular PDx list and the second (G70) being a “catch-all” ADRG for the remainder of episodes from the *Diseases & Disorders of the Digestive System* MDC.

Regarding Figure 22, the PDx A04.7 in the ADRG G67 is associated with an additional \$2,500 to \$4,000 across each of the groups. For the 443 cases where A04.7 occurs as a PDx with no additional diagnosis, A04.7 adds approximately an extra \$2,500 compared to the average PDx's in this DRG. Where A04.7 (352 cases) occurs with 4 other diagnoses, A04.7 adds approximately an extra \$4,000 compared to the average episode with 5 diagnoses in this DRG. The extra costs for the PDx is maintained irrespective of the number of additional diagnoses associated with it.



Figure 20: Surgical DRG - PDx with a cost at $i=1$ of more than 3 times greater than the average diagnosis within ADRG F14.

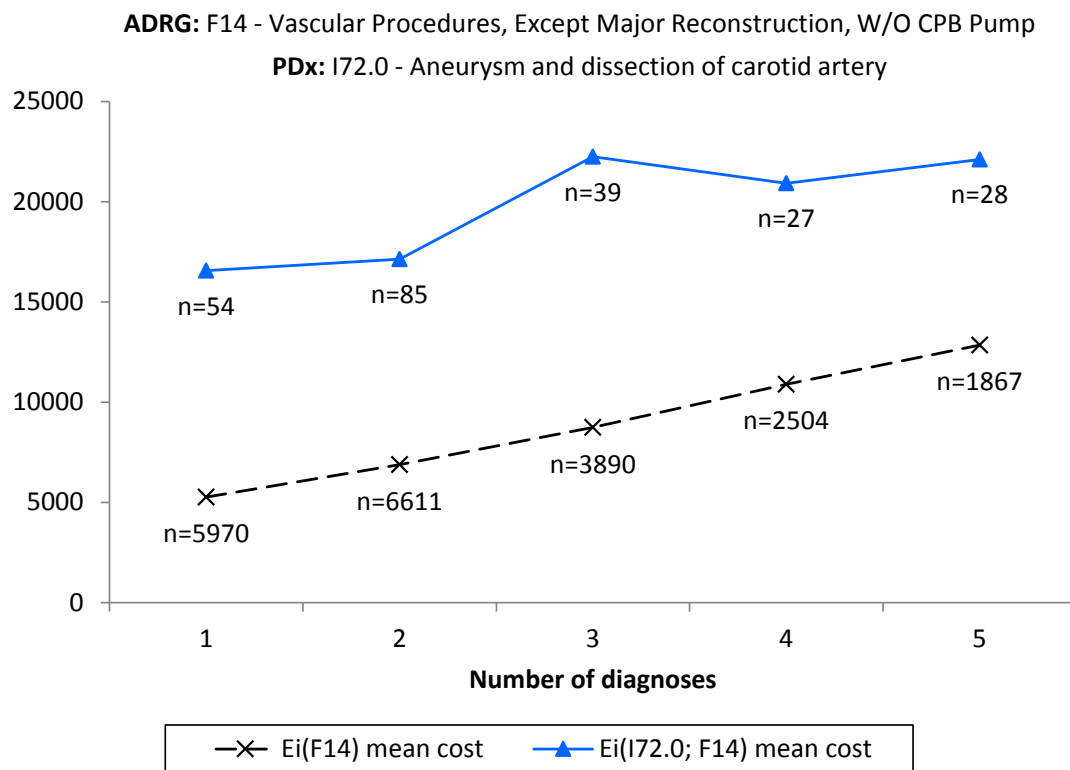


Figure 21: Surgical DRG - PDx with a cost at $i=1$ twice the average diagnosis within ADRG I20.

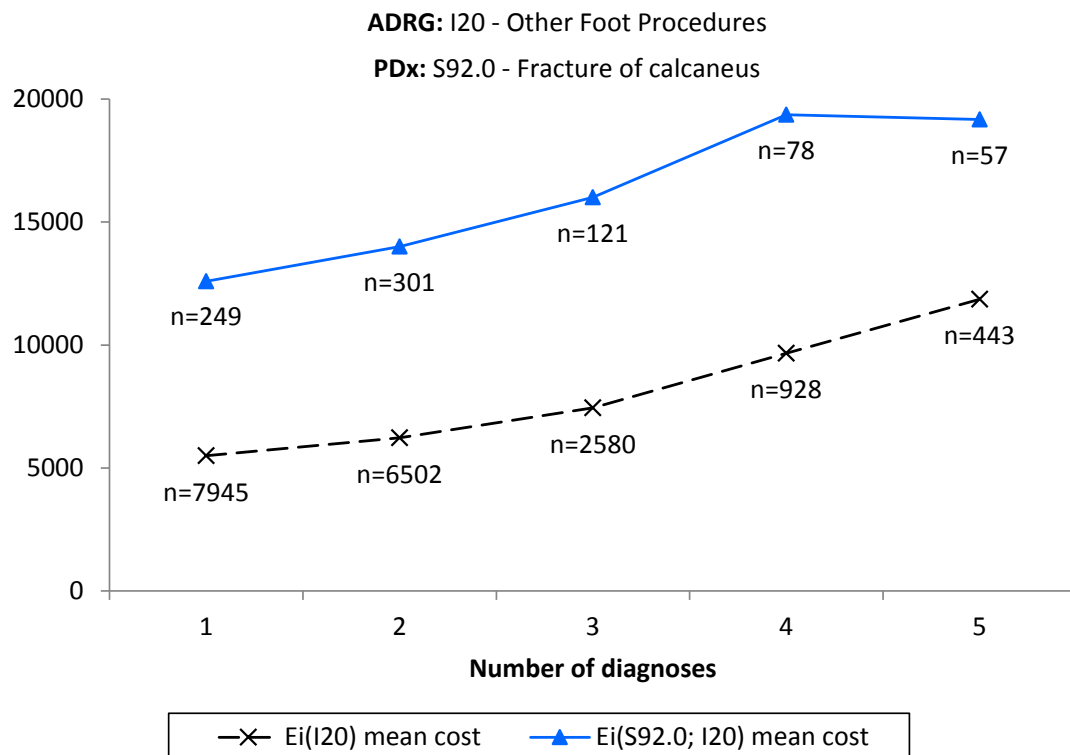




Figure 22: Medical ADRG - where a specific PDx adds additional resources.

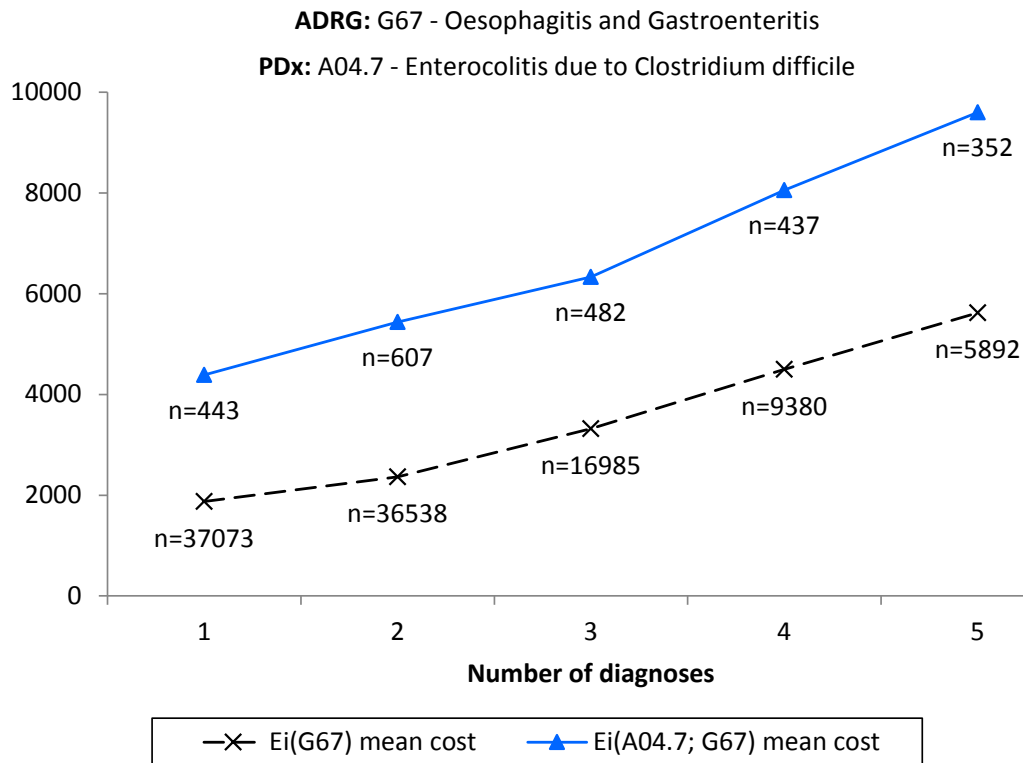
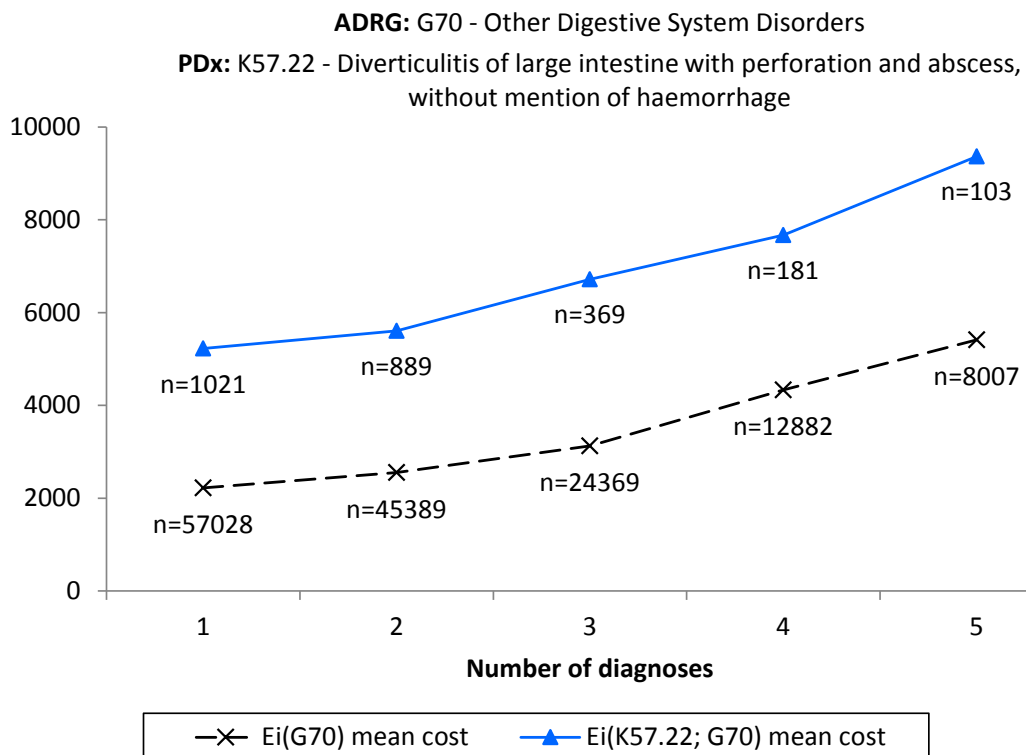


Figure 23: Other Medical DRG - contains the remainder of PDxs that are not assigned to a specific medical ADRG within the MDC.





Regarding Figure 23, the PDx of K57.22 in ADRG G70 adds \$3,002 to \$3,951 across each of the groups. For the 1,021 cases where K57.22 occurs as a PDx with no ADx, K57.22 adds approximately an extra \$3,002 compared to the average PDxs in this ADRG. Where K57.22 (103 cases) occurs with 4 other diagnoses, K57.22 adds approximately an extra \$4,000 compared to the average episode with 5 diagnoses in this ADRG. The extra costs for the PDx is maintained irrespective of the number of ADx associated with it.

By including PDx in the grouping process, the ECC Model is considering all available patient clinical information to determine DCLs and an ECCS. Conversely, ignoring the PDx once the ADRG is selected discards important information, notably for surgical ADRGs but also for some medical ADRGs.

The proposed DCL process does not distinguish between PDx and ADx in assessing the impact of diagnoses on cost variation within ADRGs. In doing so, the ECC Model does not use the terms 'complication' and 'comorbidity' as they are no longer an accurate description of the information used in calculating complexity and do not adequately represent the proposed case complexity adjustment process.

Key Finding 5

Unlike the existing system, the PDx has been recognised to contain important information on complexity over and above its use in allocating an episode to an ADRG. This is especially true for surgical ADRGs. Many principal diagnoses are in fact combinations of health conditions, and include vital information on conditions that would otherwise be additional diagnoses.

Recommendation 3

Based on Key Finding 5, given the noted degree of variation in episode cost according to the PDx, ACCD in consultation with the DTG and CCAG recommends that the PDx be included in the construction of DCLs, reflecting the information contained in many principal diagnoses on the complexity of a case within its assigned ADRG.



6 Guiding principles for Diagnosis Complexity Level assignment

The identification of diagnoses permitted to be assigned nonzero DCLs (i.e. a score of 1, 2, 3, 4 or 5) has been guided by principles that aim to characterise the scope of the ECC Model in terms of diagnoses considered relevant for DRG classification purposes. The diagnoses identified as out of scope are called exclusions, some of which are excluded unconditionally and others excluded conditionally based upon certain criterion. This section presents the guiding principles and the resulting list of unconditional and conditional code exclusions for AR-DRG V8.0 and future AR-DRG versions

In developing the methodological approach to unconditional and conditional exclusions, empirical data was initially interrogated by a team of ACCD Classification Specialists followed by consideration of identified unconditional and conditional exclusions using a small subgroup of Health Information Managers from the DTG. The resultant list of unconditional and conditional exclusions was then reviewed by the CCAG for input and finalisation.

Clinical determination of conditional exclusions (CEs) for all (approximately) 16,000 codes was not possible within the timeframe. Simply accepting the current exclusions would have only resulted in exclusions on clinical grounds for a minority of designated comorbidities.

In many cases diagnosis codes are routinely used in association with other codes. To prevent 'double counting', an algorithm has been developed to ensure that these diagnosis codes are removed. This method ensures that the proposed exclusion process can be operationalised and provide stability over time. Maintaining the existing (since AR-DRG V4.0) code exclusions without having a reason to do so was not feasible.

The following guiding principles identify the scope of the DCL within the ECC model and provide a fundamental underpinning to case complexity processing for AR-DRG V8.0.



6.1 Diagnosis Complexity Level Assignment

The first stage of the ECC model takes an acute admitted episode and assigns a DCL to each of the episode's allocated diagnosis codes. These DCLs measure the relative level of cost associated with the diagnoses and are specific to the ADRG to which the episode belongs.

The DCLs of all possible combinations of diagnosis codes and ADRGs are stored in the DCL matrix/array, although many combinations may never occur in practice. The DCL scope of the ECC model is defined as those diagnosis codes that may receive a nonzero DCL in at least one ADRG; diagnoses falling outside the DCL scope (i.e. diagnosis codes that are always assigned a DCL of zero regardless of the ADRG in which they occur) are called unconditional exclusions (UEs). The following guiding principles have been used to identify diagnosis codes that may be assigned a nonzero DCL.

For purposes of defining DCL scope within the ECC model (i.e. defining the unconditional and conditional exclusions), ICD-10-AM diagnosis codes have been divided into the following four groups.

- Group 1: Chapter 18 Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified (R00-R99)
- Group 2: Chapter 21 Factors influencing health status and contact with health services (Z00-Z99)
- Group 3: Unacceptable principal diagnoses
- Group 4: Special case exclusions

Care was taken to exclude diagnoses that would not normally be assigned in an acute admitted setting and/or could undermine the ECC Model by either providing perverse incentives for code assignment and/or allowing the model to be devalued by the indiscriminate assignment of codes for unspecified, ill-defined and transient conditions.



6.2 DCL scope guiding principle 1

To assist in defining the DCL scope, the following guiding principle has been used to identify excluded codes, both unconditional (or UEs) and conditional (CEs), which have been assigned a DCL of zero.

DCL scope guiding principle 1 excludes codes that provide additional or supplementary information to another code already assigned. Generally, these codes:

- (a) Are of ill-defined and/or transient conditions or symptoms that may be best classified to other more specific chapters within the classification,
- (b) Provide context rather than information critical to the clinical description of an acute admitted episode of care, or
- (c) Identify a characteristic that is already captured by other diagnosis codes present on the record of the acute admitted episode of care.

Applying DCL guiding principle 1, UE codes have been broken into the following four groups. There are exceptions among Groups 1 to 4, which remain within the DCL scope (i.e. may receive a nonzero DCL), and these have been identified using DCL guiding principle 2, detailed below.

6.2.1 Group 1: Chapter 18 Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified (R00-R99)

Group 1 codes (see Appendix 2) were considered UEs as they are generally unspecified, ill-defined and/or transient conditions that may be best classified to other chapters within the classification.



6.2.2 Group 2: Chapter 21 Factors influencing health status and contact with health services (Z00-Z99)

Group 2 or Z codes (see Appendix 3) were all considered UEs as they provide context and are in themselves not acute conditions but are used:

'for occasions when circumstances other than a disease, injury or external cause classifiable to categories A00–Y89 are recorded as 'diagnoses' or 'problems'. This can arise in two main ways:

- *When a person who may or may not be sick encounters the health services for some specific purpose, such as to receive limited care or service for a current condition, to donate an organ or tissue, to receive prophylactic vaccination or to discuss a problem which is in itself not a disease or injury.*
- *When some circumstance or problem is present which influences the person's health status but is not in itself a current illness or injury. Such factors may be elicited during population surveys, when the person may or may not be currently sick, or be recorded as an additional factor to be borne in mind when the person is receiving care for some illness or injury.'* (ICCC, 2012b, p.501)

6.2.3 Group 3: Unacceptable principal diagnoses

Codes that can only appear as an ADx (i.e. codes that are considered unacceptable principal diagnoses, as per AR-DRG 961Z) (see Appendix 4). External cause, place of occurrence, activity and morphology codes have also been included in this group.



6.2.4 Group 4: Special case exclusions

Special case exclusions include the following sub-groups:

Codes that add descriptive information to an already assigned ICD-10-AM code

These codes have been considered UEs and include:

- *Bacterial, viral and other infectious agents* (B95 – B97). These codes are assigned as additional diagnoses to a condition classified in another ICD-10-AM chapter to add descriptive information (when available) about the microorganism causing the condition.
- *Delivery* (O80 – O84). In most instances, these codes are assigned as a principal or additional diagnosis with a condition classified elsewhere in Chapter 15, or to add descriptive information about the mode of delivery.
- *Fetus and newborn affected by maternal factors and by complications of pregnancy, labour and delivery* (P00 – P04) and *P96.4 Termination of pregnancy, affecting fetus and newborn*. These codes are assigned as additional diagnoses to a condition classified in another ICD-10-AM chapter to add descriptive information about the cause of the condition.
- *Burns involving less than 10% of body surface, with less than 10 % or unspecified full thickness burns* (T31.00). This UE is supplementary to categories T20-T25 and T29 which provide greater detail in relation to complexity from within the 3-character category according to erythema, partial thickness or full thickness burns of specific body sites
- *Diabetes mellitus and impaired glucose with multiple microvascular complications* (E09 – E14 with 4th and 5th character subdivision of .71). These codes have been identified as cluster codes that always accompany the microvascular complication diabetes codes assigned from the E09 – E14 code range.

Sequelae (late effect) codes not appearing in Group 3.

Unacceptable principal diagnoses are not directly related to the immediate clinical profile of the patient but to one of causality and have been considered UEs (See Appendix 5).



Full-time dagger (aetiology) codes.

These codes are always accompanied by asterisk (manifestation) codes which, as a consequence of the dagger code exclusion, will be assigned the joint DCL of the dagger/asterisk pair. These dagger codes have been considered to be UEs (See Appendix 6).

Conditional Exclusions (CEs)

These codes have been identified for particular dagger asterisk pairs of DCL in-scope codes. In these cases, the dagger code is excluded from being assigned a DCL whenever the asterisk code is present (i.e. asterisk code excludes dagger code from receiving a DCL).

A dagger asterisk pair¹⁴ (x,y), where x is the dagger code and y is the asterisk code, is defined as a CE whenever y is a full-time asterisk code and x is the only dagger code associated with y. In this case, (y,x) has been defined as a CE, where y excludes x from receiving a DCL. Appendix 7 lists the CEs.

6.3 DCL scope guiding principle 2

6.3.1 Codes identified as 'in scope' for DCL assignment from Groups 1– 3 above

The DCL guiding principle 2¹⁵ has been used to identify ICD-10-AM diagnosis codes from Groups 1 -3 that are defined as within the DCL scope (i.e. capable of being assigned a nonzero DCL). That is, these have been identified as exclusions from the unconditional exclusion list.

DCL scope guiding principle 2 identifies codes from groups 1 - 3 above that are capable of providing information critical to the clinical description of an acute admitted episode of care. The following in-scope codes from Group 1 to 3 have been identified for inclusion:

¹⁴ (x†,y*)

¹⁵ DCL scope guiding principle 2 identifies codes from groups 1 - 3 above that are capable of providing information critical to the clinical description of an acute admitted episode of care.



- Codes from Chapter 18 Signs and symptoms (from Group 1) have been determined in scope for a DCL as they meet ACS 1802 Signs and symptoms, specifically point f: “certain symptoms, for which supplementary information is provided, that represent important problems in medical care in their own right.” Codes include:
 - R02 *Gangrene, not elsewhere classified*
 - R15 *Faecal incontinence*
 - R18 *Ascites*
 - R32 *Unspecified urinary incontinence*
 - R40.2 *Coma, unspecified*
 - R56.- *Convulsions, not elsewhere classified*
 - R57.- *Shock, not elsewhere classified*
 - R65.- *Systemic inflammatory response syndrome (SIRS)*
 - R95.- *Sudden Infant death syndrome*
- *Open wound* categories: S01.-, S11.-, S21.-, S31.-, S41.-, S51.-, S61.-, S71.-, S81.- and S91.- (from Group 3)
- S06.05 *Loss of consciousness of prolonged duration (more than 24 hours) without return to pre-existing conscious level* (from Group 3)
- Group 3 codes from within the code range T31.10 – T31.99 *Burns classified according to extent of body surface involved*. NB: Category T31.00 is a UE (see point 4 in 6.2.4 above, p. 82)
- Codes from Chapter 21 *Factors influencing health status and contact with health services* including:
 - Z06 *Resistance to antimicrobial drugs* (appear in Group 2 and 3)
 - Z21 *Asymptomatic human immunodeficiency virus (HIV) infection status*
 - Code range Z34 to Z35 *Supervision of normal and high risk pregnancies*
 - Z92.1 *Personal history of long term (current) use of anticoagulants*

Total number of UEs identified: 4,285 codes (not counting morphology codes). Most (3,105) are external cause, place of occurrence and activity codes.



DCL assignment: Of the total 16,708 diagnosis codes (not counting morphology codes), 12,423 codes remain in-scope for DCL assignment within the ECC Model.

Key Finding 6

The list of diagnoses permitted to be assigned nonzero DCLs has been guided by principles that aim to characterise the scope of the ECC model in terms of diagnoses considered relevant for DRG classification purposes. The diagnoses identified as out of scope are called exclusions, some of which have been excluded unconditionally and others excluded conditionally (depending on other diagnoses present) based upon guiding principles for DCL assignment.

Recommendation 4

Based on Key Finding 6, ACCD in consultation with the DTG and CCAG recommends that the proposed guiding principles for DCL assignment and list of identified exclusions (unconditional and conditional) be adopted.



7 The use of condition onset flag in classification design

As per Australian Coding Standard 0048 'The condition onset flag (COF) is a means of differentiating those conditions which arise during, from those arising before, an admitted patient episode of care.' Permissible values are:

- COF 1: A condition which arises during the episode of admitted patient care and would not have been present or suspected on admission.
- COF 2: A condition previously existing or suspected on admission such as the presenting problem, a comorbidity or chronic disease.

This section considers whether the COF should have an impact on case complexity when the condition arose during the acute admitted episode (COF = 1). Specifically, the purpose of evaluating the COF was to ascertain whether conditions with a COF of 1 should be excluded from the case complexity adjustment process in classification design.

7.1 Background¹⁶

Early case complexity research using manual chart review revealed that much patient variation in complexity or cost was driven by 'adverse patient occurrences' rather than underlying poor health on admission (Schumacher, et al., 1987). Over the past decade, increasing policy attention has focused on the quality and safety of acute admitted care. This has led to a number of approaches to distinguish comorbidities (those conditions present on admission) from complications (conditions arising during the episode of admitted patient care). The fundamental argument behind making this distinction is that hospitals should not be 'rewarded' when poorer quality care leads to higher costs.

In 2008, the US Medicare program adopted a 'Hospital Acquired Conditions' (HAC) policy (CMS, 2008) that removed 10-15 HACs from CC lists of MS DRGs used for hospital payment. These conditions were chosen because they 'could reasonably have been prevented

¹⁶ This background is a summary of the discussion paper: *What approaches to DRG development have been taken internationally to account for complications and comorbidities?* Prepared by Associate Professor Terri Jackson with assistance by ACCD staff for purposes of the AR-DRG Classification System Development and Refinement project (see Appendix 1)



through the application of evidence-based guidelines'. Some of them (e.g. in-hospital falls) are applied across all DRGs, while others (e.g. 'surgical site infections following certain orthopedic procedures') are removed from a more limited set of intervention-specific records before DRG assignment. The payment effect of this policy has been relatively small, but the cultural shifts attributed to the policy were larger than might have been expected (McNair, et al., 2009b).

Information to identify HACs had been recorded for some time in Canada, and in Victoria. A 'present on admission' flag was subsequently adopted in California and New York state, with a national Medicare mandate to record this data element adopted to support the HAC policy within months of the Australian agreement amongst the States to record such a marker. The Australian data element is termed the condition onset flag (COF), and like others, distinguishes between comorbidities recorded as being present on admission and those noted in the patient's record as arising during the admitted episode of care.

Australian research using the COF in Victorian data estimated that 15.7 per cent of hospital expenditures were attributable to additional costs in cases with hospital-acquired diagnoses (HADs) (Ehsani, et al., 2006). Later work on data from two states (Victoria and Queensland) found they added 17.3 per cent to treatment costs (Jackson, et al., 2011). While not all HAD complications can be prevented with current medical knowledge, these studies demonstrate the need to assess the merits of excluding HADs from the AR DRG classification.

Building on this work, McNair, et al., (2009a) used Victorian data to model a payment system that would result in hospitals being paid a 'complications-averaged' DRG amount by removing COF diagnoses from DRG grouping and from estimation of relative resource weights. Modelling showed that while 15 per cent of episodes contained a HAD, only 1.6 per cent of episodes were grouped to a different DRG. If such a payment policy were applied to a sample of metropolitan hospitals, the model found that payments would be redistributed from +1.8 per cent to -2.5 per cent per hospital.

A joint working party (JWP) of IHPA and the Australian Commission on Safety and Quality in Health Care have published two literature reviews on options for integrating quality incentives into hospital pricing (JWP, 2013). More recently the JWP has commissioned work



to estimate the effect of HACs on hospital costs, but results of this analysis have not yet been made publicly available.

7.2 Is there a potential role for COF in classification development?

In considering the role of conditions arising during the episode of care and their importance in classification development it is useful to identify the differences between five separate scenarios:-

Scenario 1: A patient is assigned to a specific ADRG. The patient has a PDx of “XXX.XX” and no ADx.

Scenario 2: A patient has a PDx of “XXX.XX” and a single Adx of “YYY.YY”. The ADx was present on admission (COF=2).

Scenario 3: A patient has a PDx of “XXX.XX” and a single ADx of “YYY.YY”. The ADx was not present on admission (COF=1) and arose during the episode of care by chance, perhaps due to the patient’s overall medical condition.

Scenario 4: A patient has a PDx of “XXX.XX” and a single ADx of “YYY.YY”. The ADx was not present on admission (COF=1) and arose during the episode of care due to hospital actions.

Scenario 5: A patient has a PDx of “XXX.XX” and a single ADx of “YYY.YY”. The ADx was not present on admission (COF=1) and arose during the episode of care due to hospital actions but the patient was known to be at extreme risk of having the complication. This is a special sub-group of patients with scenario 4.

These five scenarios are summarised in Table 10 below.

**Table 10: Characteristics of patients in five scenarios.**

Condition	Senario 1	Senario 2	Senario 3	Senario 4	Senario 5
Adx Present on the record	No	Yes	Yes	Yes	Yes
Adx Present arising during admission	No	No	Yes	Yes	Yes
Adx caused by hospital	No	No	No	Yes	Yes
High Risk of Adverse Event during Treatment	No	No	No	No	Yes

Different classification design issues require analysis of different combinations of patients with these different scenarios. One possible approach would be to exclude codes that relate to hospital error. However the COF is not sufficient for this task as we cannot distinguish between patients in scenarios 3 (arising spontaneously) and scenario 4 (arising due to hospital actions). Excluding ICD codes with a COF of 1 will inappropriately remove ADx for patients under scenario 3.

Furthermore, the question arises as to whether a condition with a COF of 1 is caused by the acts or omissions of the hospital, or merely associated with the hospital episode (for example, all falls in a person with known high falls risk cannot be entirely prevented). A response by hospitals to total exclusion risks the exclusion of patients in Scenario 5 from treatment.

This issue of funders not paying for hospital error was recognised by McNair et al (2009) who argued that by retaining patients in scenarios 3 and 4 within the “uncomplicated” DRG when calculating cost weights, all hospitals would be funded for an underlying proportion of error that might reasonably relate to patients in scenario 3 (ADx spontaneously arising). This argument however still relies on the appropriateness of the classification structure that recognises the impact of ADx that are present on admission (COF=2).

Excluding codes with a COF of 1 prior to these analyses would result in patients in scenario 4 (and 5) being treated as not having the ADx, driving up the costs of the baseline and distorting the DCL. In our terms, removing these codes would shift patients with higher



costs from E_i to E_{i-1} and distort the calculation of mean $(E_{i,ADx}) \div \text{mean}(E_{i-1})$. This in turn could affect the splitting of ADRGs.

Excluding ICD codes with a COF of 1 from the development of the classification might reduce the capacity for the classification to accurately capture cost differences and thereby reduce the capacity of pricing policy to make appropriate adjustments for funding poor quality care.

7.3 Other Considerations

7.3.1 Retaining the integrity of the COF information

Excluding diagnoses with a COF of 1 which resulted in lower funding (either through classification design or pricing policy) could result in hospitals failing to identify the cases appropriately within the clinical record. Therefore the COF would need to be part of any clinical audit program.

7.3.2 Treating at risk patients

Some treatments, by their very nature, have inherent risks of misadventure. Again excluding diagnoses which resulted in funding which was significantly under cost could result in selection bias in those patients offered specific therapies. High risk patients who are likely to undergo misadventure and require unfunded care might be less likely to be offered treatment. This would only become an issue where the COF resulted in lower funding.

7.4 In summary

The classification should reflect, as far as reasonable, the real cost of treating patients. Excluding codes with a COF of 1 has the potential to distort these comparisons through shifting relative high cost cases to 'uncomplicated' groups. This in turn could reduce the potential for pricing policy to make appropriate adjustments to not reward poor quality care.



Key Finding 7

In considering the potential role of the condition onset flag (COF) within the classification, ACCD had difficulty in defining what a condition arising during the episode of care meant in terms of its preventability. It was determined that removing codes associated with conditions arising during the episode of care (COF = 1) from the complexity algorithm would reduce the capacity of the classification to explain true cost differences between DRGs. It would potentially alter incentives to treat patients with risks of complication.

Recommendation 5

Based on Key Finding 7, ACCD in consultation with the DTG and CCAG recommends that the COF should not be used to exclude diagnosis codes from the DRG development process.



8 Evaluation of performance of the ECC Model

PCCL is used as an ADRG splitting variable in the AR-DRG V7.0 classification. The episodes of each ADRG are divided into up to five categories by PCCL, and these set of categories are used together with other episode characteristics to define the DRGs of each ADRG.

Throughout this section, the splitting of ADRGs into their PCCL categories is referred to as the PCCL model. This section focuses on performance tests undertaken using 5-category ECCS models and makes comparisons of these models' performance against that of the PCCL model and the AR-DRG classification.

The comparative performance of models was assessed using two R-squared indices for the goodness of fit. Both indices take values from 0 (a model that does not describe data at all) to 1 (a model that fits all data perfectly).

The first index is the classical R-squared statistic (R^2). This index is used in analysis of variance in linear modelling and measures the fraction of variance in data that is explained by the model. Sometimes it is called the reduction in variance (RIV) due to the model.

The second index belongs to the category of pseudo- R^2 statistics and is sometimes called the Kullback-Leibler R^2 (Cameron & Windmeijer, 1996; 1997). This index results from the analysis of deviance in generalised linear modelling. Deviance plays a similar role in generalised linear modelling as does variance in linear modelling, thus the Kullback-Leibler R^2 is referred to here as the reduction in deviance (RID) due to the model.

To enable some form of comparative assessment of performance between PCCL and ECCS, the episodes of each ADRG were partitioned into up to five categories based on their ECCS. This was done by determining all possible partitions of each ADRG into a maximum of five ECCS-based categories where each category had at least 5 per cent and 100 of the ADRG's episodes.

The set of all possible ECCS partitions were then tested for their ability to explain costs using the two model fit statistics R^2 and RID. That is, two 5-category ECCS models were defined, one which is optimised by R^2 and the other optimised by RID. Although the two approaches



(i.e. R^2 and RID optimised) lead to very similar models and results, both approaches are presented to allow full consideration of each.

The following comparative performance tests were undertaken using these 5-category ECCS models:

1. Comparison against the PCCL model;
2. Comparison against the AR-DRG classification;
3. Performance on paediatrics episodes; and
4. Performance on geriatrics episodes.

The ECCS models perform exceptionally well compared to the partitioning of ADRGs by PCCL and even compared to the full DRG classification. Table 11 provides a comparative summary of the performance of each model in terms of R^2 and RID statistics.

Table 11 shows that the ECCS-based models outperform both the PCCL model and the AR-DRG classification in terms of R^2 , but the AR-DRG classification performs slightly better than the ECCS models in terms of RID.

The AR-DRG classification performs better overall when degrees of freedom (i.e. numbers of categories) are taken into account, with 766 AR-DRG categories present in the data compared to 1,744 for the ECCS-based models. However, as will be seen in the following sections, the higher performance of the AR-DRG classification is driven by the use of LOS as an ADRG splitting variable within the classification (e.g. same-day DRG splits). This observation, together with the ECCS models' significantly higher performance compared to the PCCL model, demonstrate that if PCCL were replaced by ECCS, the resulting classification would significantly outperform AR-DRG V7.0. The performance of the ECC Model as an episode complexity measure within the AR-DRG classification will be the subject of Phase Two of AR-DRG V8.0 development.

**Table 11: Comparative summary of the performance of ADRG splitting models.**

Classification	Categories	R ²	RID
ADRG	402	42.8%	60.9%
PCCL (in ADRG)	1,908	48.8%	67.7%
DRG	766	52.0%	70.7%
5-category ECCS model - R ² optimised	1,743	55.1%	69.8%
5-category ECCS model - RID optimised	1,744	54.9%	69.9%

The DRG classification has a maximum of four categories by ADRG, and although the ECCS and PCCL models each have a maximum of five categories, the count of categories within each ADRG often varies considerably between the models.

Table 12 summarises the percentage split of episodes among the five categories of each model. DRGs are allocated to Categories 1 to 4 based on the number of DRGs in the ADRG and the corresponding DRG levels. Specifically:

- all Z-level DRGs are allocated to Category 1;
- for ADRGs with two DRGs, the B-level DRG is allocated to Category 1 and the A-level DRG is allocated to Category 2;
- for ADRGs with three DRGs, the C-level DRG is allocated to Category 1, the level-B DRG is allocated to Category 2 and the A-level DRG is allocated to Category 3; and
- for ADRGs with four DRGs, the D-level DRG is allocated to Category 1, the C-level DRG is allocated to Category 2, the B-level DRG is allocated to Category 3 and the A-level DRG is allocated to Category 4.

Note that the second category of the PCCL model (i.e. PCCL = 1) only contains 0.4 per cent of all episodes. From this perspective, it would appear PCCL is more similar to a four-category classification. However, Table 11 shows that the PCCL model contains 160 more categories than either of the two ECCS models – an average of 4.7 PCCL categories per ADRG compared to an average of 4.3 ECCS categories per ADRG. The main factor reducing the number of



ECCS-based categories is the requirement that each contain at least 5 per cent and 100 of an ADRG's episodes.

Table 12: Summary of percentage distribution of episodes by PCCL, DRG and ECCS categories.

Category	PCCL	DRG	ECCS - RID optimised	ECCS - R ² optimised
Category 1	77.0%	77.1%	57.7%	60.4%
Category 2	0.4%	18.5%	21.7%	21.2%
Category 3	8.7%	4.2%	10.3%	9.1%
Category 4	8.3%	0.2%	6.5%	5.6%
Category 5	5.6%	0.0%	3.8%	3.6%

Table 13 shows the distribution of ADRGs based on the number of categories each has for the four models. It shows that the DRG classification has by far the most ADRGs with a single category (i.e. not split), followed by the ECCS models, with all but two ADRGs having three or more PCCL categories. Instances of fewer than five ECCS categories are a consequence of the minimum sample criteria applied to each ECCS category.

Table 13: Breakdown of ADRGs within each model by number of model categories.

Classification	One Category	Two Categories	Three Categories	Four Categories	Five Categories
PCCL	1	1	19	61	321
DRG	128	192	77	6	0
ECCS - RID optimised	18	20	42	54	269
ECCS - R ² optimised	18	20	42	55	268

Note that this breakdown is based on categories appearing in the data and may not represent the total number of categories theoretically possible. The factor that most distinguishes performance of ECCS models compared to the PCCL model at the ADRG level is the minimum 5 per cent and 100 episodes property that the ECCS models satisfy. No such minimum exists for PCCL, with 385 of the 1,908 PCCL categories containing less than 1 per cent of their respective ADRG episodes and 163 categories containing less than 10 episodes.

On the other hand, the factor that most distinguishes performance of ECCS models compared to the AR-DRG classification is the use of LOS as a splitting variable in the AR-DRG classification (i.e. <5 days, <2 days or same-day). Regarding minimum episodes by DRG, only eight DRGs contain less than 5 per cent of their respective ADRGs, and only two DRGs contain less than 100 episodes. These comparisons are explored further in the following sections.

8.1 Comparison of the 5-category ECCS models against the PCCL model

The performance of the ECCS models was compared at the ADRG level by looking at the difference in either of the RID or R^2 performance statistics. A positive difference in RID between the models (e.g. RID of ECCS model minus RID of PCCL model) indicates better performance of the ECCS model, and a negative difference indicates better performance of the PCCL model. Figure 24 shows the distribution of these RID differences across all ADRGs, with the equal performance threshold of zero highlighted red.

Figure 24: Comparative difference of RID statistics across ADRGs – ECCS models v PCCL model.

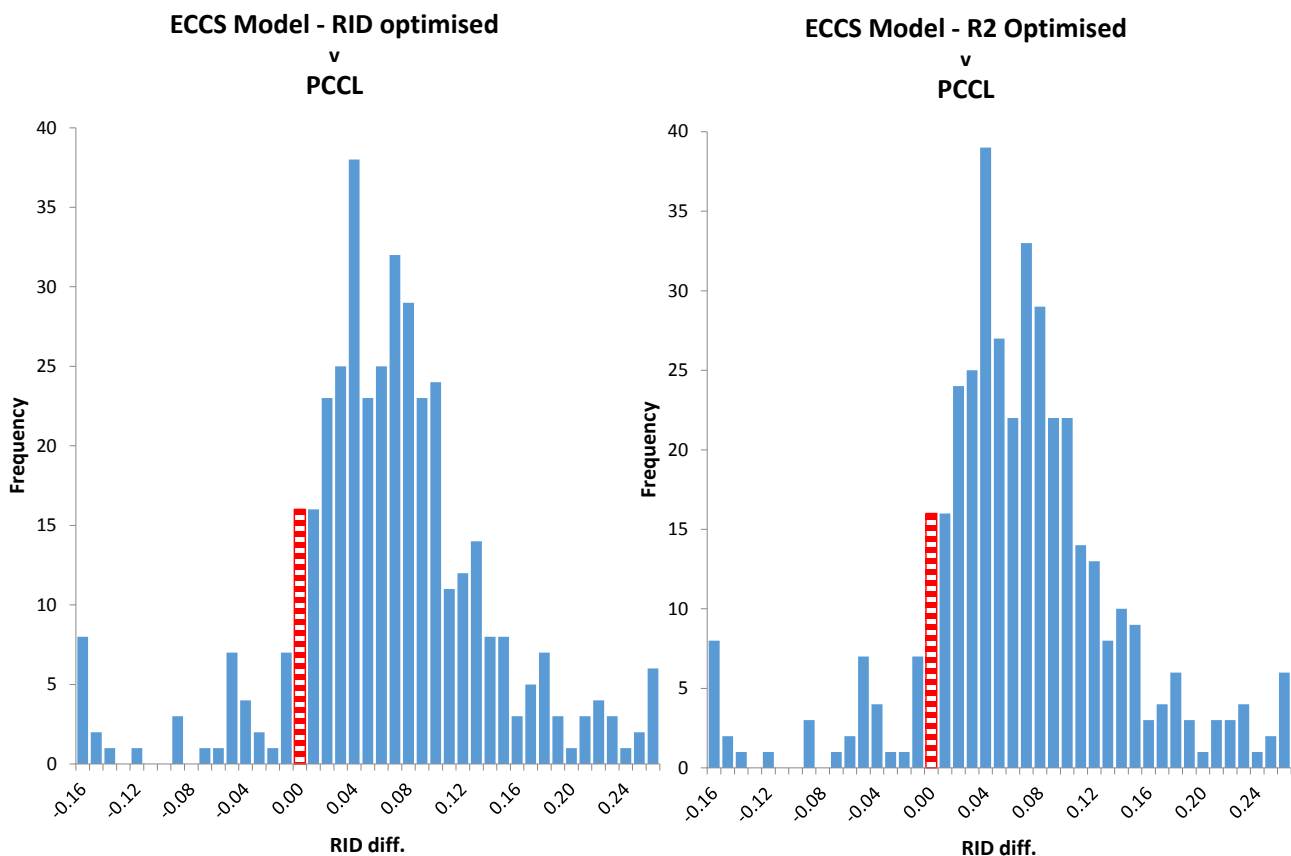
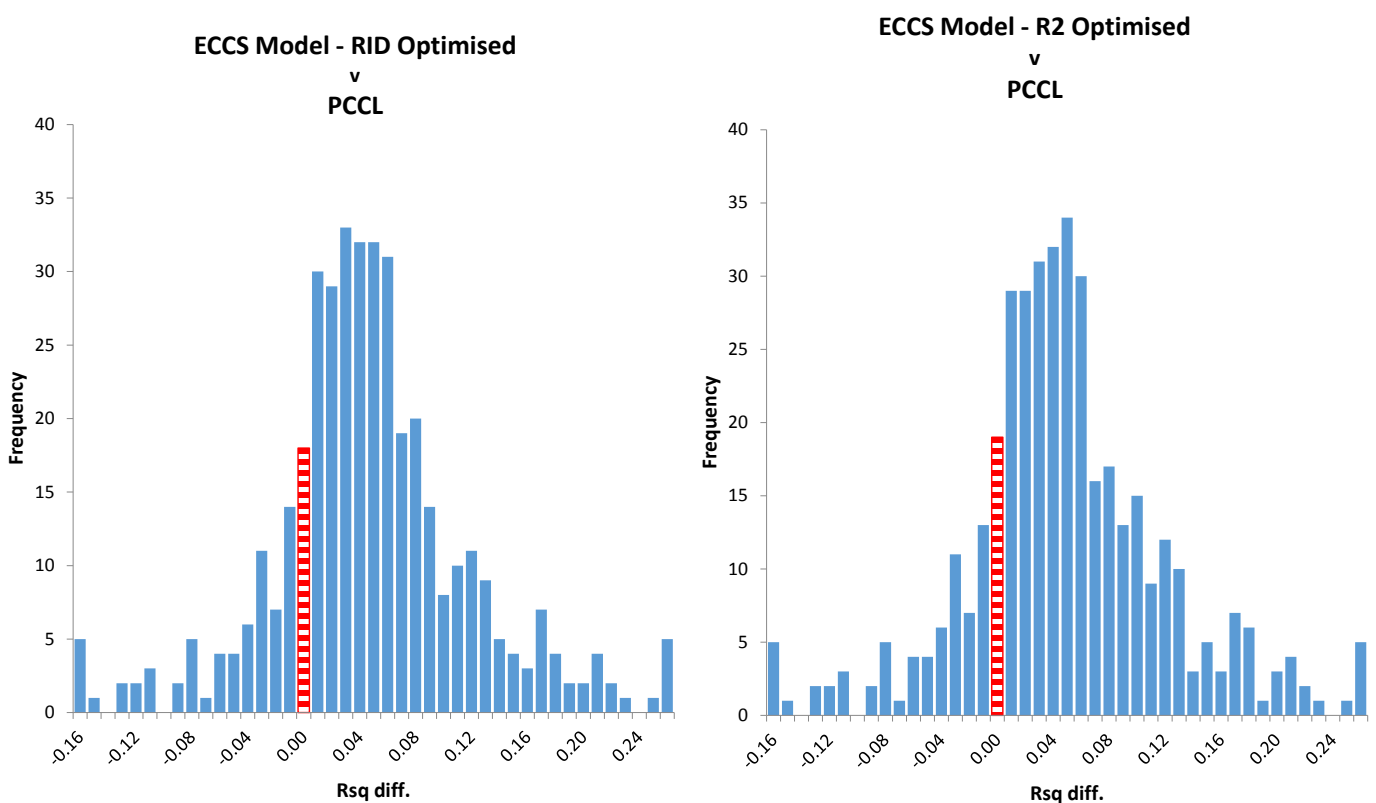


Figure 24 shows that the two 5-category ECCS models perform significantly better than the PCCL model across almost all ADRGs, based on comparison of RID. Specifically, the ECCS models outperform the PCCL model in 358 of the 403 ADRGs. Furthermore, of the 45 ADRGs where the PCCL model performs at least as well as the ECCS models, almost half of them are due to the ECCS models only having one category. Given that the minimum sample criteria (i.e. 5 per cent and 100 episodes) is the main cause of some ADRGs only having one ECCS category, relaxing these criteria would lead to the ECCS models performing better than PCCL in almost all (if not all) ADRGs.

Similar to Figure 24, Figure 25 shows the comparative difference between R^2 statistics of models across ADRGs. As with the RID comparison, Figure 25 shows the ECCS models demonstrate significantly better R^2 performance across almost all ADRGs.

Figure 25: Comparative difference of R^2 statistics across ADRGs – ECCS models v PCCL model.



8.2 Comparison of the 5-category ECCS models against the AR-DRG classification

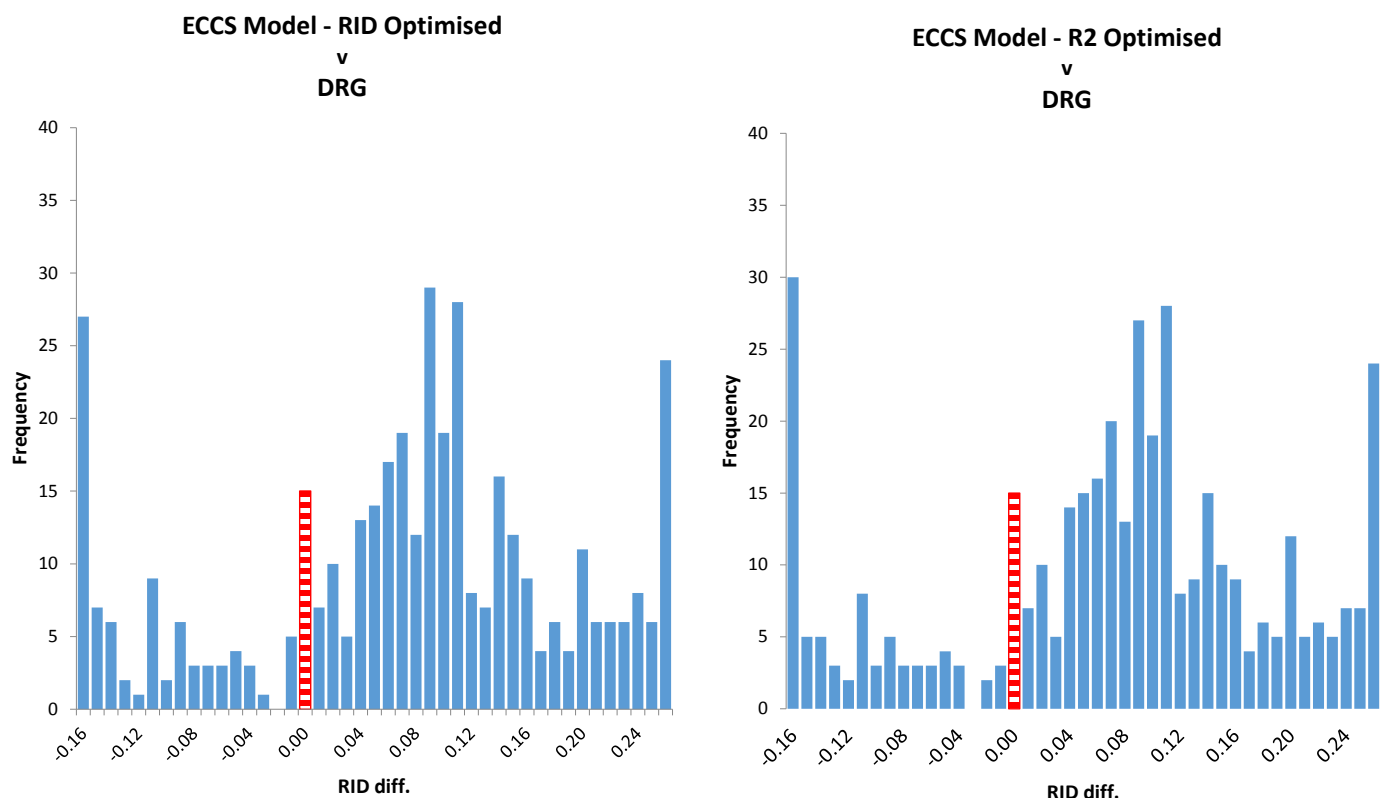
classification

In contrast to the 5-category ECCS models and the PCCL model, the maximum number of ADRG categories within the AR-DRG classification is four.

Similar to the previous section, comparative performance of the 5-category ECCS models against the AR-DRG classification is assessed by comparing R^2 and RID statistics by ADRG.

Figure 26 shows the comparative difference in the RID statistic by ADRG for the 5-category ECCS models compared to the AR-DRG classification. Although the DRG classification performs better overall in terms of RID (see Table 11), the two ECCS models perform consistently better by ADRG.

Figure 26: Comparative difference of RID statistics across ADRGs – ECCS models v DRG classification.



As indicated previously, the main factor influencing better performance of the DRG classification is the use of LOS as a splitting variable. The greater level of performance is

demonstrated in Figure 27, which is Figure 26 restricted to LOS-split ADRGs. Note that the restriction to LOS-split ADRGs excludes Z-level ADRGs that are defined using LOS.

Comparison of these two figures shows that the ADRGs in which the DRG classification is outperforming the ECCS models are made up almost entirely of ADRGs with LOS as a splitting variable. That is, the ECCS models outperform the AR-DRG classification in almost all ADRGs where the AR-DRG classification does not use LOS as a splitting variable.

Figure 27: Comparative difference of RID statistics across LOS-split ADRGs – ECCS models v DRG classification.

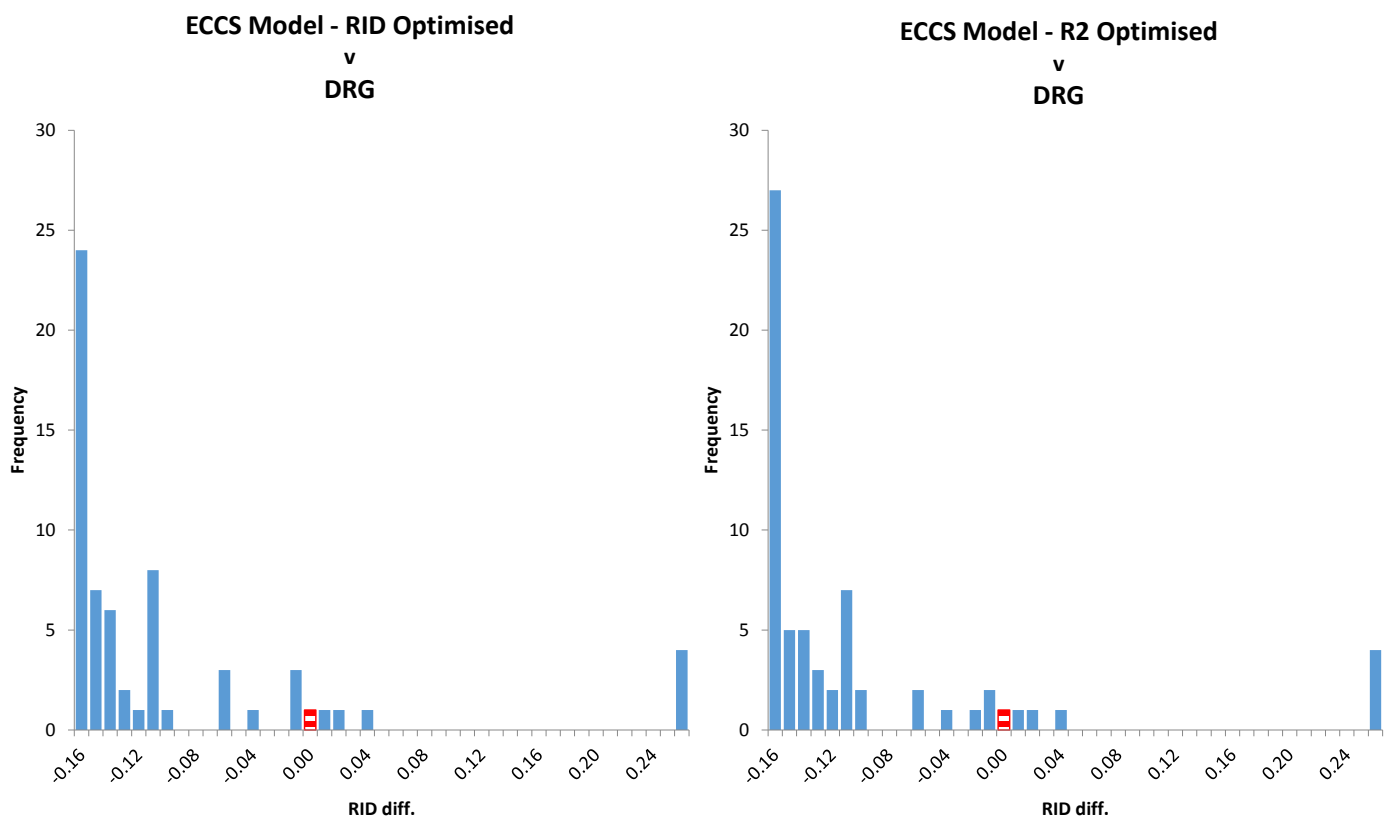
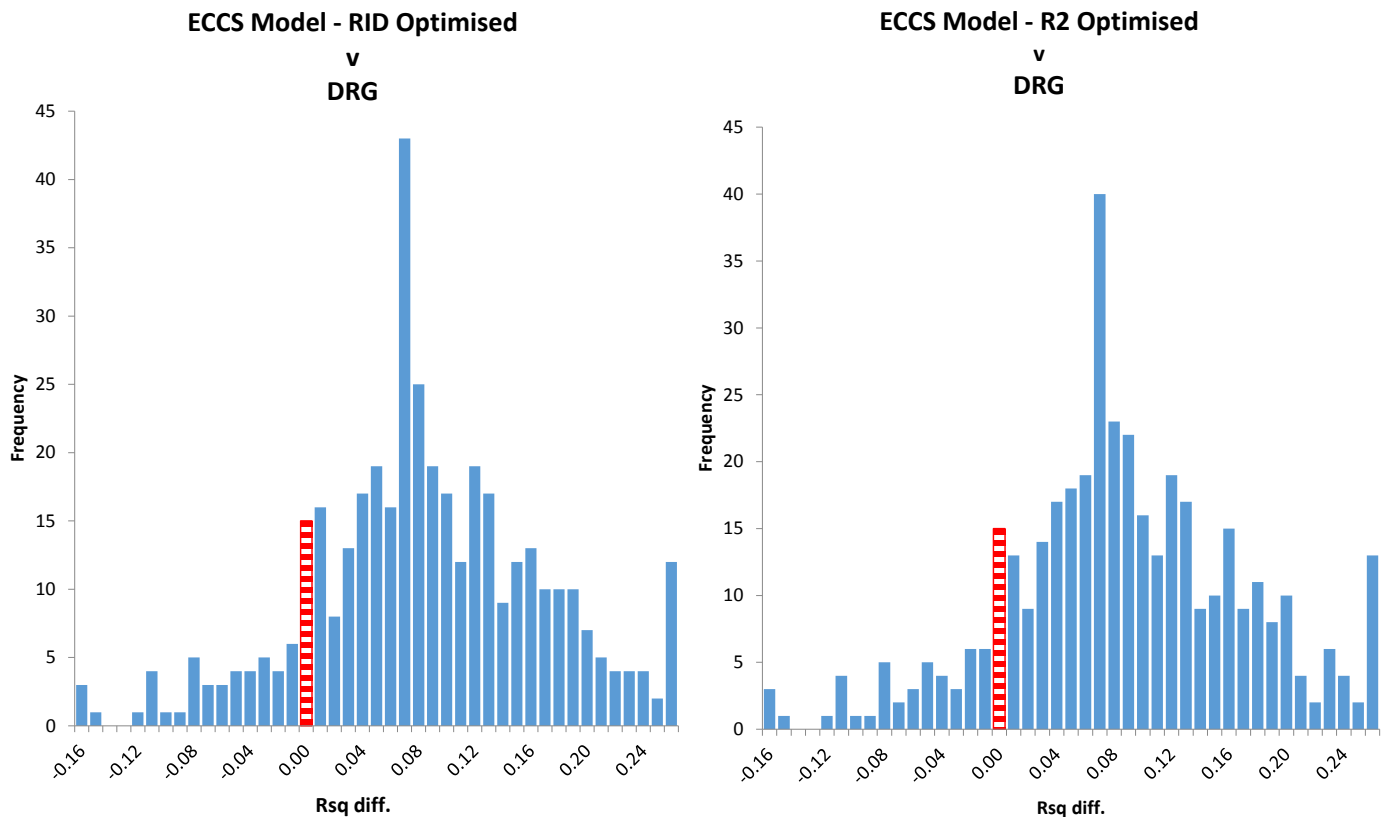


Figure 28 shows the comparative difference in the R^2 statistic by ADRG for the 5-category ECCS models compared to the AR-DRG classification. Table 11 shows that the two ECCS models exhibit significantly better overall R^2 performance compared to the DRG classification, and Figure 28 demonstrates this better performance is consistent across almost all ADRGs.

Figure 28: Comparative difference of R^2 statistics across ADRGs – ECCS models v DRG classification.



8.3 Comparison of performance across all models

Figure 29, Figure 30, Figure 31 and Figure 32 provide comparisons of the ECCS model (rounded to the nearest integer), the 5-category ECCS model (R^2 optimised), the existing PCCL model and the AR-DRG classification itself. The figures are representative of the pattern of evidence across almost all ADRGs of the ability of ECCS to outperform PCCL and the AR-DRG classification in terms of prediction of costs. In particular, comparisons of the y axis (i.e. cost) scale and range for the ECCS-based models to those of the PCCL and AR-DRG models show a significantly greater ability of the ECCS-based models to link episode complexity and cost.

Figure 29: Comparison of ECCS, PCCL and AR-DRG by cost - B02 Cranial Procedures.

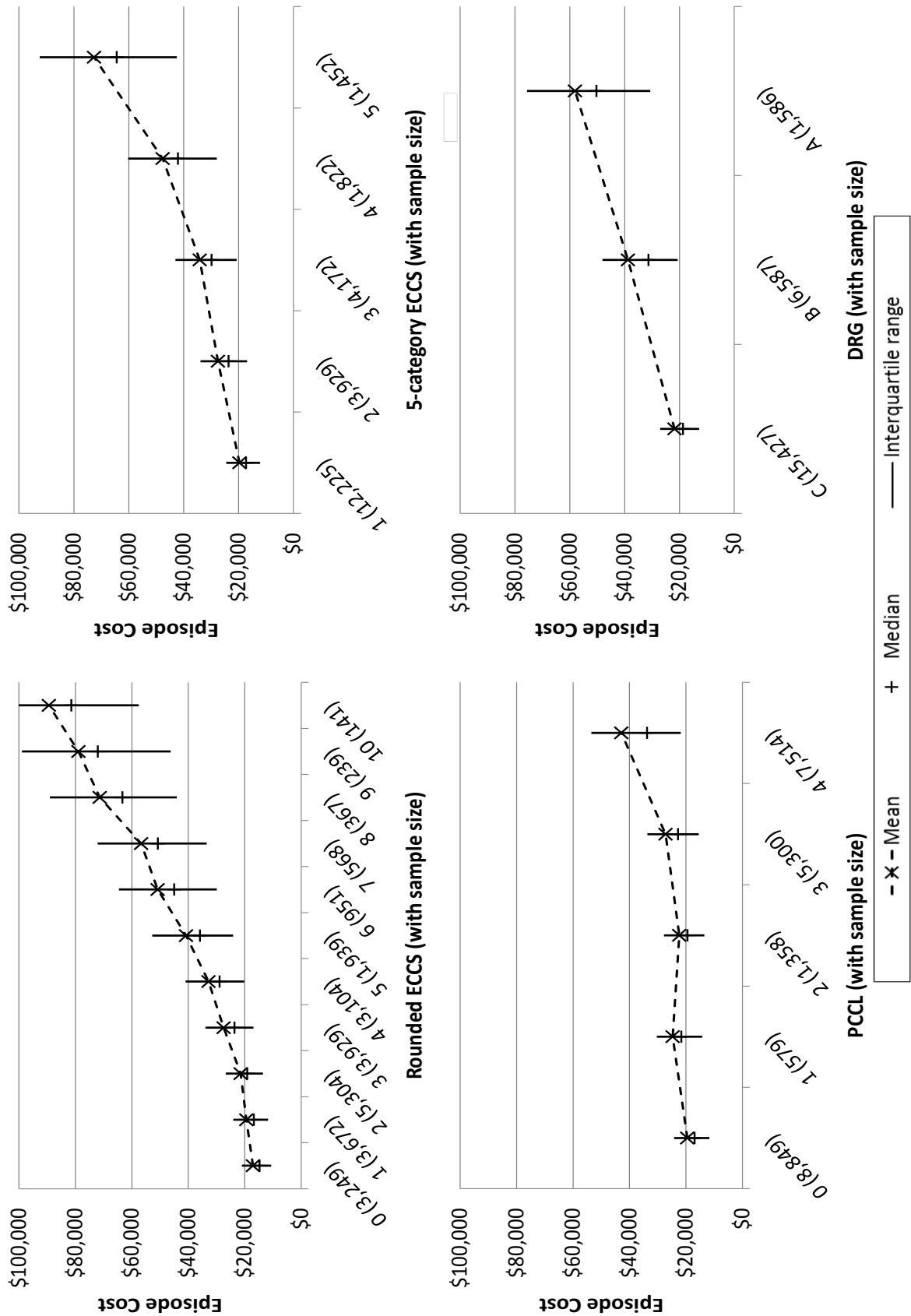




Figure 30: Comparison of ECCS, PCCL and AR-DRG by cost - G03 Stomach, Oesophageal and Duodenal Procedures.

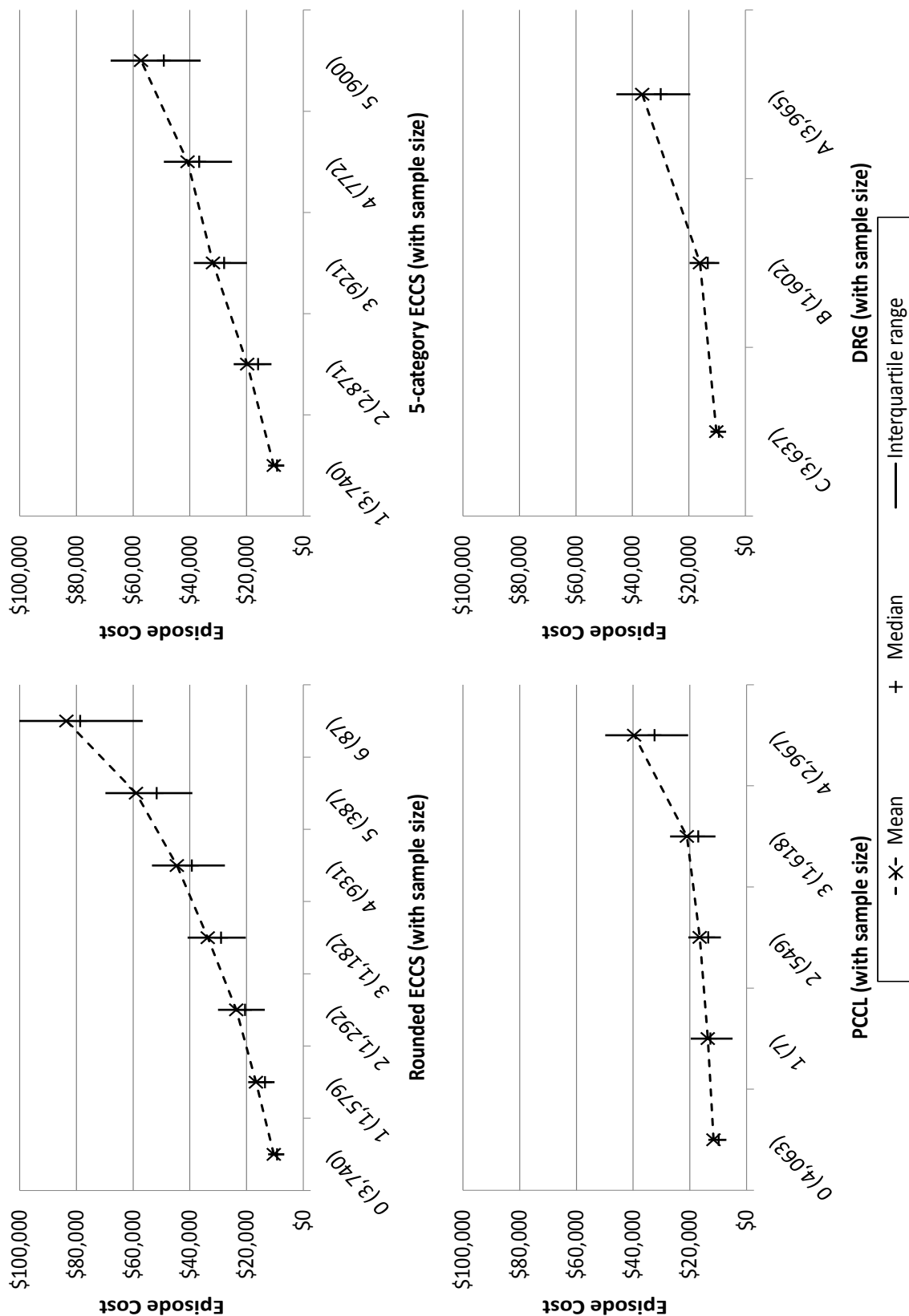




Figure 31: Comparison of ECCS, PCCL and AR-DRG by cost - H01 Pancreas, Liver and Shunt Procedures.

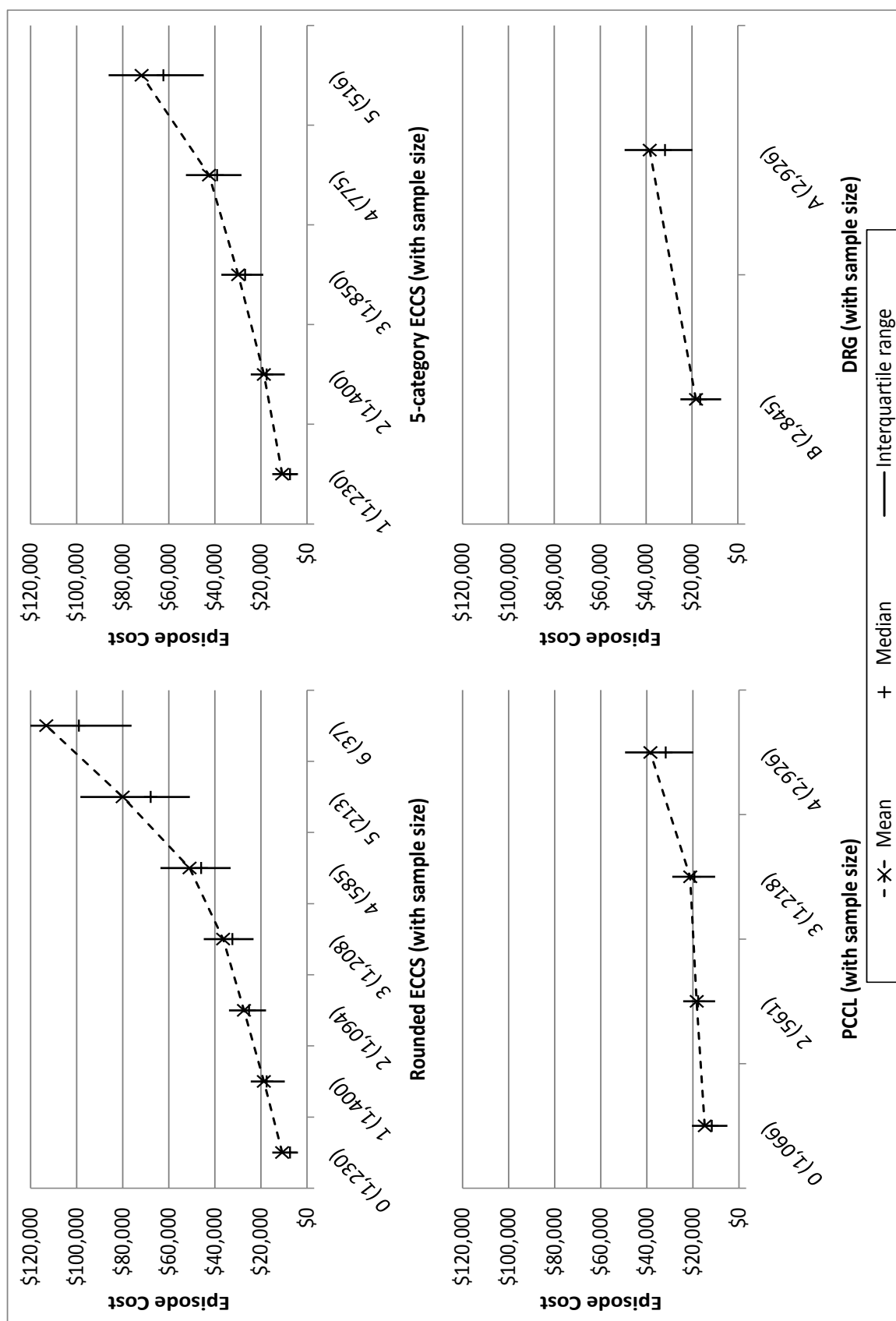
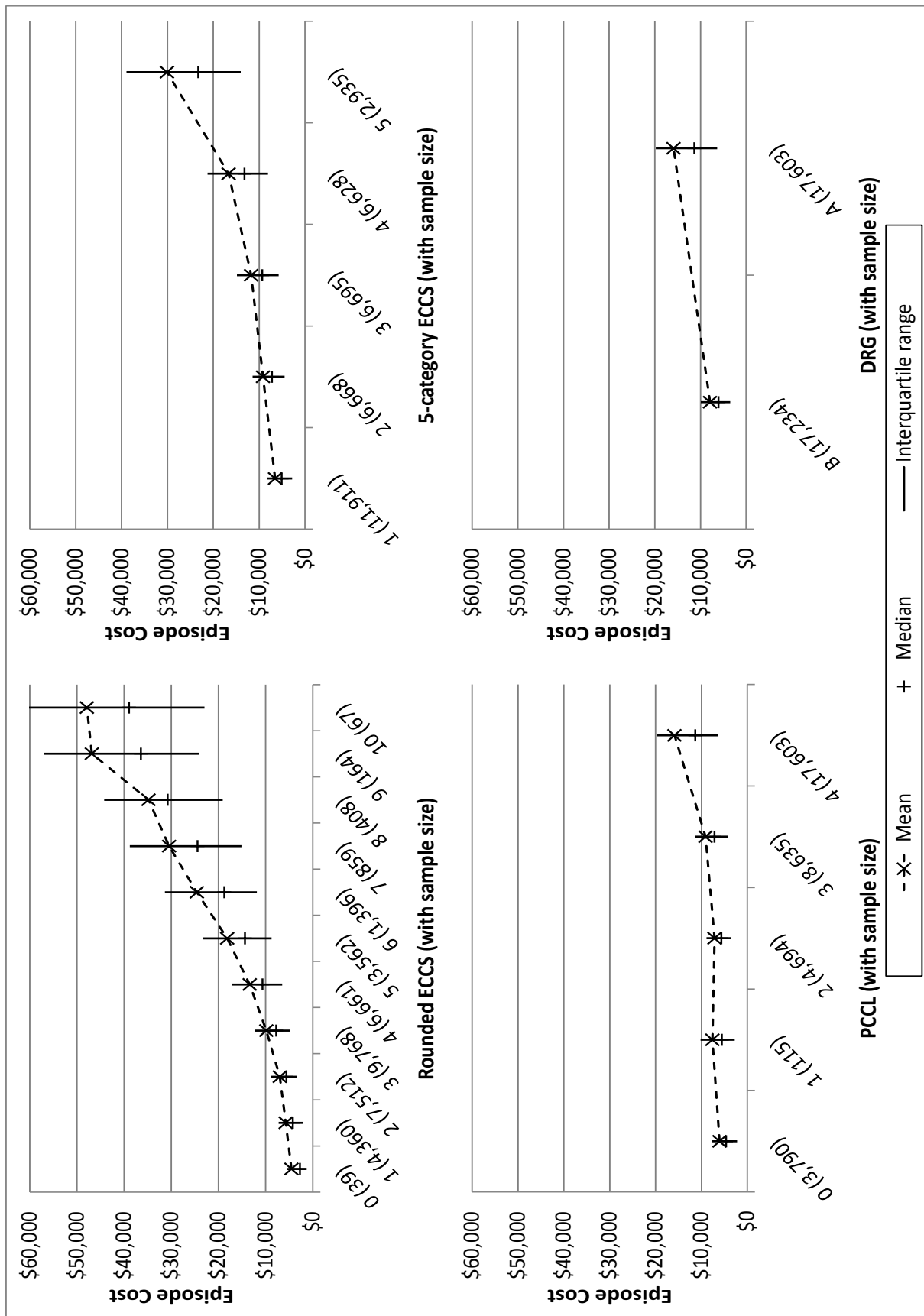




Figure 32: Comparison of ECCS, PCCL and AR-DRG by cost - T60 Septicaemia.





Key Finding 8

The new ECCS was shown to be a much improved predictor of cost at the episode level when compared to the current PCCL. Overall, the ECCS was shown to have the potential to greatly increase performance of the AR-DRG classification.

Recommendation 6

Based on Key Finding 8, ACCD in consultation with the DTG and CCAG recommends that the ECCS measure be adopted to estimate clinical complexity at the episode level.

Recommendation 7

Based on Key Findings 1 – 8, ACCD in consultation with the DTG and CCAG recommends that the proposed ECC Model which has shown to be a much improved predictor of cost at the diagnosis and episode level be adopted as the new case complexity structure for AR-DRG Version 8.0 and future versions of the AR-DRG classification.

8.4 Comparative performance on paediatric episodes

The assessment of the ECC Model on paediatric episodes was assessed in terms of the overall RID and R^2 statistics on paediatric episodes and a cost ratio analysis of population-level cost models restricted to paediatric episodes.

Note that paediatrics was taken to be episodes with age (in years) of patient at admission less than or equal to 16, excluding newborn episodes (i.e. MDC 15).

Table 14 shows that the 5-category ECCS models perform well overall in terms of R^2 and RID statistics. Similar to performance across all episodes, the ECCS models significantly outperform both the nested PCCL model and the AR-DRG classification in terms of R^2 . They also outperform the PCCL model in terms of RID, and although they have a lower RID compared to the AR-DRG classification, the extra reduction in deviance of the AR-DRG classification is due to the use of other variables including LOS in the classification.



Table 14: Comparative summary of the performance of ADRG splitting models using R² and RID - paediatric episodes.

Classification	Categories	R ²	RID
ADRG	375	37.8%	45.1%
PCCL (in ADRG)	1,626	44.7%	52.2%
DRG	701	44.9%	56.6%
ECCS 5Cat (in ADRG) - R ² optimised	1,563	49.5%	54.9%
ECCS 5Cat (in ADRG) - RID optimised	1,568	48.7%	55.0%

Next, the ability of each model to predict costs was examined in terms of cost ratios.

Specifically, a cost model was derived for each of the four models by taking the mean cost of each model's categories across the entire population of episodes. That is, with reference to the category counts in Table 11, the mean cost was taken across each of the 1,908 PCCL categories, across each of the 766 DRGs, across each of the 1,743 ECCS (R² optimised) categories, and across each of the 1,744 ECCS (RID optimised) categories. This resulted in four mean-cost models.

These mean-cost models were then applied across all data, giving a model predicted value against each episode for each of the four models. The data was then restricted to paediatric episodes, and the ratio of total actual costs to total model costs was then calculated for each model.

The resulting cost ratio of each model can be thought of as the adjustment required to re-calibrate the model to the paediatric subpopulation, with a cost ratio above 1 indicating that the given model under-costs paediatric episodes.

The closer the paediatric cost ratios are to the value of 1, the less calibration the model would require, and hence the better the model performs with respect to the paediatric episodes in the sense that it shows less bias to cost estimation.

Table 15 shows the cost ratios of each of the four models when applied to the paediatric episodes. The ECCS models are seen to demonstrate significantly better performance in terms of minimising bias of cost estimation, showing an adjustment of 0.7 per cent would be

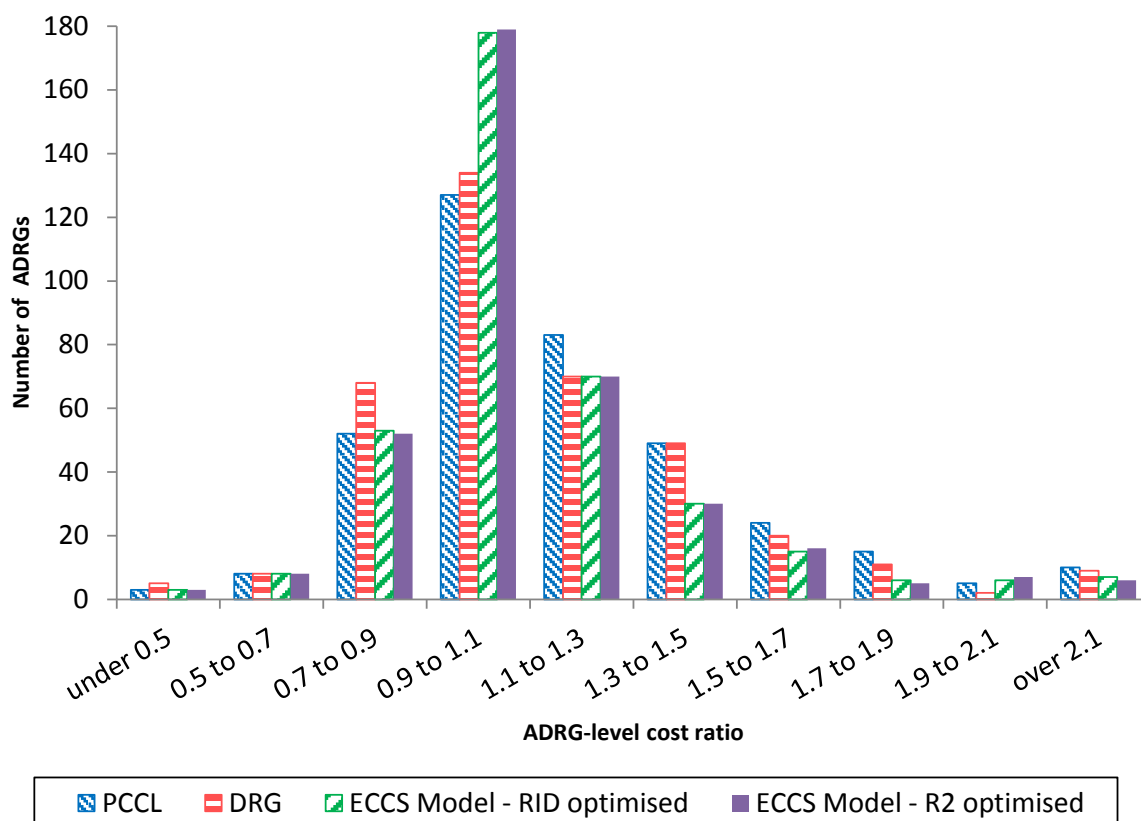
required to calibrate the model to paediatric episodes, compared to 3.4 per cent for the DRG model and 6.6 per cent for the PCCL model.

Table 15: Comparison of cost ratios on paediatric episodes.

Classification	Cost Ratio
PCCL (in ADRG)	1.066
DRG	1.034
ECCS 5Cat (in ADRG) - R ² optimised	1.007
ECCS 5Cat (in ADRG) - RID optimised	1.007

The cost ratio of each ADRG was calculated in the same way and used to measure the model performance on paediatric episodes at the ADRG level. Figure 33 shows the distribution of ADRG-level cost ratios for each of the four models. The two ECCS models show a significantly higher proportion of ADRG-level cost ratios between 0.9 and 1.1.

Figure 33: Distribution ADRG-level cost ratios across each model on paediatric episodes.





In summary, compared to both PCCL and the AR-DRG classification, the ECCS-based models exhibit a significantly enhanced ability to classify paediatric episodes in a way that minimises bias in cost estimation. The decreased bias in cost estimation occurs consistently within ADRGs and also manifests in an overall decreased bias in cost estimation.

8.5 Comparative performance on geriatric episodes

Similar to the evaluation of performance on paediatric episodes, the models were compared in terms of their performance of the subset of geriatric episodes. For the purpose of this analysis, geriatric episodes were defined as those with age (in years) of patient at admission of greater than or equal to 80.

Table 16 summarises the overall performance of the ECCS-based models when restricted to geriatric episodes, comparing them to ADRGs, the nested PCCL model and the AR-DRG classification.

As with their performance overall and when restricted to paediatric episodes, the ECCS-based models have significantly better R^2 performance on geriatric episodes compared with the other models. The ECCS models also perform significantly better than the nested PCCL model in terms of RID, and have only a slightly lesser performance than the full AR-DRG classification in terms of RID. As discussed with regard to overall performance, the slightly increased performance of the full AR-DRG classification is due to the use of non-clinical variables including LOS in the classification.

Table 16: Comparative summary of the performance of ADRG splitting models using R^2 and RID - geriatric episodes.

Classification	Categories	R^2	RID
ADRG	364	38.7%	60.3%
PCCL (in ADRG)	1,539	46.7%	68.5%
DRG	689	48.9%	70.9%
ECCS 5Cat (in ADRG) - R^2 optimised	1,578	51.4%	70.2%
ECCS 5Cat (in ADRG) - RID	1,579	51.2%	70.3%



Classification	Categories	R ²	RID
optimised			

Table 17 shows the cost ratios of each model applied to geriatric episodes. While the ECCS models do not perform as well on these episodes as they do on paediatric episodes, their geriatric cost ratio of 1.015 (or 1.5 per cent increase) is relatively minor and does not differ substantially from that of the DRG and PCCL models.

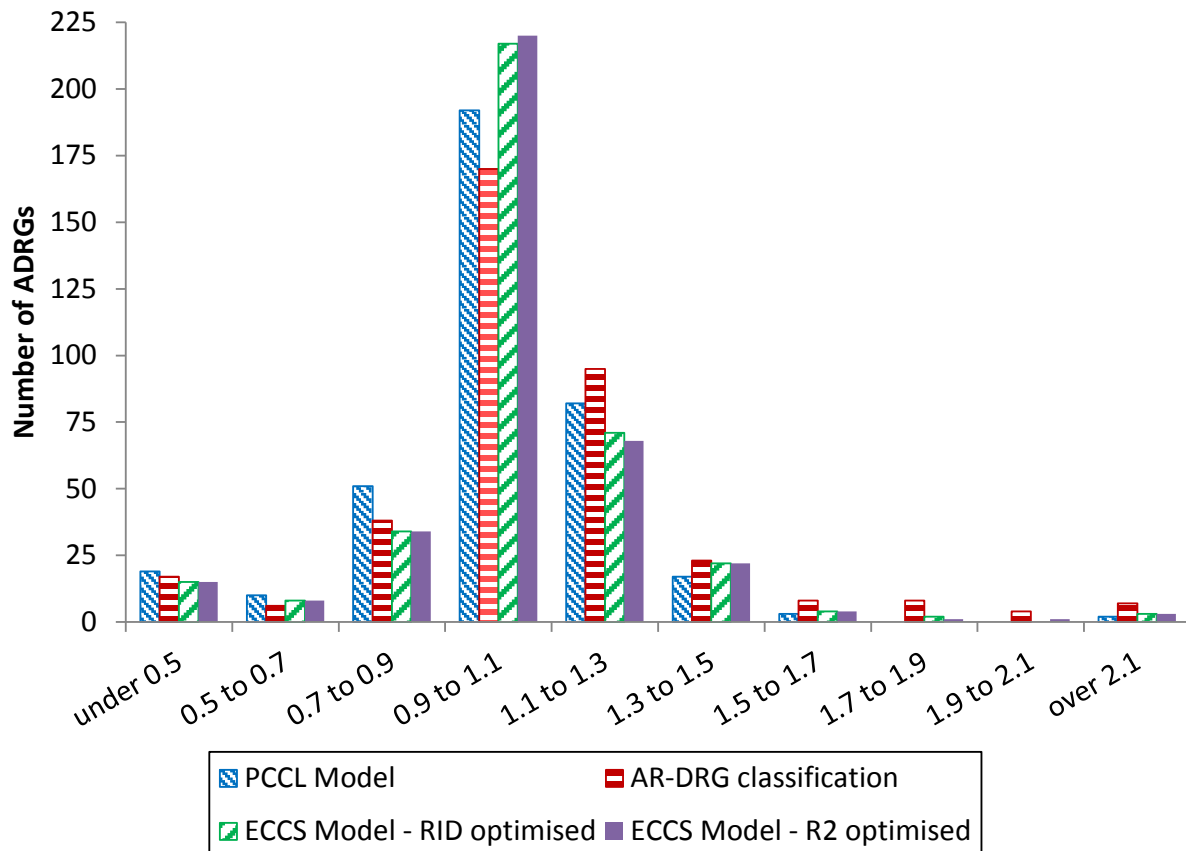
Among the contributing factors to the 1.5 per cent under-prediction of costs across geriatric episodes is the ECC Model scope excluding diagnoses associated with long-stay episodes such as those where the patient is awaiting admission to a residential age care service (Z75.11).

Table 17: Comparison of cost ratios on geriatric episodes.

Classification	Cost Ratio
PCCL (in ADRG)	0.991
DRG	1.007
ECCS 5Cat (in ADRG) - R ² optimised	1.015
ECCS 5Cat (in ADRG) - RID optimised	1.015

Finally, the performance of the four models was evaluated by calculating their cost ratios restricted to geriatric episodes within each ADRG. Figure 34 shows the distributions of these ADRG-level cost ratios for each of the four models. Similar to their performance on paediatric episodes, Figure 34 shows a considerably tighter distribution of the ECCS models' ADRG cost ratios about 1, compared to the PCCL model and AR-DRG classification.

Figure 34: Distribution ADRG-level cost ratios across each model on geriatric episodes.



In summary, compared to the PCCL model and the AR-DRG classification, the ECCS models show an enhanced capability to minimise bias in cost estimation within ADRGs. Although there does remain more of an overall bias in cost estimation of geriatric episodes compared to performance of the ECCS models on paediatric episodes, this is understandable given the increased presence of non-clinical factors that may lead to increased length of stay, such as the processes of admission to residential aged care facilities.

Key Finding 9

ECCS performance was evaluated on paediatric and geriatric episodes and compared to that of the current PCCL measure. When compared to the PCCL measure, ECCS showed a much improved ability to minimise bias in cost estimation within ADRGs among both cohorts (i.e. minimising over and under prediction of cost).



Recommendation 8

Based on Key Finding 9, ACCD recommends that separate approaches for paediatric and geriatric episodes are not required, given the improved performance of the ECC Model in explaining cost variations for paediatric and geriatric episodes.



9 Continued refinement of ECC Model

The ECC Model will require ongoing evaluation and refinement to ensure the model is best suited to its role in the AR-DRG classification. This section sets out a framework that supports a continuous and systematic approach to evaluation and refinement which encompasses the broader AR-DRG refinement process.

The ongoing evaluation and refinement of the ECC Model can be divided into two groups, relating to:

- The methodological and technical components of the model, including processes that enhance model precision and performance within the AR-DRG classification; and
- The empirically derived components of the model, including processes that monitor for unintended model behaviour caused by anomalous data and ensure stable updating of the model as new data becomes available.

9.1 Ongoing evaluation and refinement of methodological and technical components of the ECC Model

The methodological and technical components of the ECC Model ensure that the model has a sound conceptual basis and that its performance is optimised against its role in the AR-DRG classification. Therefore any refinement of these components of the model should be approached in a way that maintains consistency with the underlying methodology either by extension or modification. Any refinement of this type should also be measured in terms of its ability to enhance the performance of the model within the AR-DRG classification.

One particular element of the ECC Model technical specifications that should be evaluated on an ongoing basis with a view to enhancing the model's performance is the level of precision of the DCLs. The following section details the process by which this could begin to occur.



9.1.1 Enhancing DCL precision

As detailed in the *Derivation of the Diagnosis Complexity Level* section, the finest level of precision at which the ECC Model derives DCLs is the three-character category level within CDC. This level is chosen to strike a balance between maximising both ECCS performance and robustness of DCL values whilst at the same time applying a consistent method of DCL derivation across all diagnosis codes.

The ongoing evaluation and review of the ECC Model should include careful consideration of extending this aspect of the methodology by increasing DCL precision for identified diagnosis codes. This would occur by deriving DCL values at the fourth and fifth character level of the diagnosis codes in circumstances where comparative levels of complexity between codes is known (e.g. when disease severity is explicit in the code descriptions).

The following three-character categories are identified as initial candidates for enhanced DCL precision by calculating DCL at the fourth and fifth character level:

- I70.2- *Atherosclerosis of arteries of extremities*
- I83.- *Varicose veins of lower extremities*
- L89.- *Decubitus ulcer and pressure area*
- N18.- *Chronic kidney disease*
- T31.- *Burns classified according to extent of body surface involved*

Where there is adequate sample size and where calculated DCL hierarchies align with disease severity, the DCLs should be set at the enhanced level of precision. Where there is misalignment, DCLs should be combined to ensure they do not contradict comparisons of disease severity.



9.2 Ongoing evaluation and refinement of empirically derived components of the ECC Model

The ECC Model parameters, including DCLs, are derived from the three years of cost data from 2009-10 to 2011-12. Consequently, the parameters are subject to variation in data and the model is specifically adapted to the casemix and cost profiles of those three years.

As more recent data becomes available the ECC Model would benefit from having its parameters updated to ensure that its role within the AR-DRG classification remains current. The ongoing updating of model parameters is also of benefit as the quality of cost data improves over time. However, care must be taken when updating the ECC Model parameters to ensure that changes due to sample variation, or 'noise' in the data, are minimised while at the same time allowing the model to adapt to changes in casemix and costs and improvements in data quality.

The following section explores the way in which variation in data can affect DCLs, and looks at ways to minimise this effect.

9.2.1 Sample variation and DCL stability

To assess the stability of DCLs over time and with respect to data variability, the DCL matrix was calculated from different datasets.

Initial testing was undertaken based on the following approaches:

- Derivation of the DCL matrix from the full 3 years of cost data from 2009-10 to 2011-12 (i.e. the version of the DCL matrix used in the ECC Model);
- Derivation of the DCL matrix from the full 3 years of cost data from 2008-09 to 2010-11;
- 50 derivations of the DCL matrix based on randomly selected 95 per cent subsamples of establishments from the 3 years of cost data 2009-10 to 2011-12; and
- 50 derivations of the DCL matrix based on randomly selected 95 per cent subsamples of establishments from the 3 years of cost data 2008-09 to 2010-11.



The results show that, while the DCLs show a high level of stability with respect to sample variation within either of the 3-year datasets, they show considerably more variation when compared across the two 3-year datasets.

Each derivation of the DCL matrix contains 5,005,663 DCLs that aren't manually set to zero on account of being UEs (i.e. out of scope diagnoses). Table 18 shows the variation of these particular DCLs when comparing the 2009-10 to 2011-12 full sample derivation to the 50 subsample derivations of 2009-10 to 2011-12. It shows that 94.34 per cent of all subsample-derived DCLs agree with their full sample counterparts, and 99.96 per cent agree to within ± 1 DCL.

Table 18: Variation of DCLs - 2009-10 to 2011-12 full sample vs. 50 x 2009-10 to 2011-12 random 95% subsamples.

DCL difference	%
-3	0.00%
-2	0.03%
-1	2.68%
0	94.34%
1	2.94%
2	0.01%
3	0.00%
TOTAL	100.00%

Similarly, Table 19 shows the variation of DCLs when comparing the 2008-09 to 2010-11 full sample derivation to the 50 subsample derivations of 2008-09 to 2010-11. It shows that 94.10 per cent of all subsample-derived DCLs agree with their full sample counterparts, with 99.94 per cent agreeing to within ± 1 DCL.



Table 19: Variation of DCLs - 2008-09 to 2010-11 full sample vs. 50 x 2008-09 to 2010-11 random 95% subsamples.

DCL difference	%
-3	0.01%
-2	0.02%
-1	2.77%
0	94.10%
1	3.07%
2	0.02%
3	0.01%
TOTAL	100.00%

In contrast, Table 20 compares the 2009-10 to 2011-12 and 2008-09 to 2010-11 full sample DCL derivations, where 'DCL difference' is 2009-10 to 2011-12 DCL minus 2008-09 to 2010-11 DCL. It shows that only 76.66 per cent of the DCLs agree between the two 3-year windows. However, 98.53 per cent of DCLs agree to within ± 1 DCL.

**Table 20: Variation of DCLs - 2009-10 to 2011-12 full sample vs. 2008-09 to 2010-11 full sample.**

DCL difference	%
-5	0.00%
-4	0.04%
-3	0.09%
-2	0.59%
-1	12.70%
0	76.66%
1	9.17%
2	0.59%
3	0.11%
4	0.04%
5	0.00%
TOTAL	100.00%

The above tables suggest that a ± 1 DCL buffer could be used to stabilise the DCLs over time when updating using more current data. This could be achieved by halving the difference between DCL calculations and setting each 0.5 fraction of DCL to round down (i.e. 0.5 rounds down to 0, 1.5 rounds down to 1, etc.). Table 21 shows how this method would change DCLs.



Table 21: Illustration of the method changing DCLs based on halving the difference between original and revised calculation.

DCL difference	Resulting change to DCL
-5	-2
-4	-2
-3	-1
-2	-1
-1	0
0	0
1	0
2	1
3	1
4	2
5	2

Interpreting Table 20 as the change between an existing DCL matrix and a new version derived on updated data.



Table 22 shows how the resulting DCLs would change using this approach.

Table 22: Illustration of resulting changes to DCLs if a half-difference method is used to update the DCL matrix.

Change to DCL	%
-2	0.04%
-1	0.69%
0	98.53%
1	0.71%
2	0.04%
TOTAL	100.00%

In summary, a technique of halving differences could be used to increase the stability of DCLs when updating the ECC Model as new data becomes available.

Key Finding 10

Changes in clinical care and improvements in data quality over time were identified as necessitating the ongoing evaluation and review of the ECC Model to ensure it is best suited to its proposed role in the AR-DRG classification.

Recommendation 9

Based on Key Finding 10, ACCD in consultation with the DTG and CCAG recommends that an ongoing and systematic approach be taken to evaluate and refine the ECC Model as part of the broader AR-DRG refinement process.



10 Conclusion

As part of the groundwork for the development of AR-DRG V8.0 and future AR-DRG versions, ACCD undertook a significant and timely portion of refinement within the AR-DRG classification component of its work program. Given the elapse of time since the introduction of case complexity processing in AR-DRG V4.0, a review of the system was appropriate. In some cases there have been significant changes in clinical practice (e.g. reducing LOS). Further, the availability of patient level data and associated cost information is much improved, and the computing capacity to analyse the available data is now far superior to the 1990s.

During the course of the project, ACCD worked closely with both the DTG and CCAG in developing the proposed methodology for the ECC Model. A DTG subgroup of clinical and classification experts from the ACCD and the DTG reviewed and formalised guiding principles for the scope of the DCL (including unconditional and conditional exclusions) within the ECC Model. The project evolved sequentially with a number of steps undertaken to ensure its success.

A literature review and consultative process in the initial stages revealed that detailed information on the formal (i.e. theoretical) development of diagnosis level (CCLs) and episode level (PCCL) case complexity measures was lacking. Much of the earlier work on assessing CCs was based upon the extent to which LOS was increased, rather than costs. With the increased use of same day admissions the utility of LOS in describing cost differences is reducing.

In light of the lack of consensus on how increased case complexity may be best measured, ACCD determined that the revision of the current PCCL mechanism would require a sound conceptual basis and the use of robust statistical methods to ensure confidence in the resultant improvements to the current system. These underpinning principles guided the approach taken during this review.

During the exploratory stage, a thorough evidence-based review of PCCL was undertaken drawing from an available six years of patient level cost and activity data from 2006-07 to



2011-12, with cost data from the most recent three years being used for in depth study of diagnosis (CCL) and episode level complexity (PCCL).

In the first instance, cost profiles of ADRGs and diagnoses were used to test the validity of CCLs as a predictor of cost at the diagnosis level. This exploratory stage revealed that CCLs were shown to have very little (if any) correlation with cost.

Following the exploratory stage, a conceptually based, formally derived and data driven ECC Model was developed which included a DCL that measures relative costs associated with each diagnosis in the context of each ADRG; and an ECCS that measures relative costs associated with each episode based on the DCLs of the diagnoses present in the episode.

DCLs were designed to align optimally with diagnosis cost profiles and showed significantly higher correlation with costs compared to CCLs.

The ECC Model differs substantially from the approach originally taken in AR-DRG V4.0. Firstly, ACCD developed clearly defined guiding principles to characterise the scope of the ECC Model in terms of diagnoses considered relevant for DRG classification purposes. Secondly, the assignment of DCLs to selected diagnoses including the PDx, and the methodological approach taken in identifying unconditional and conditional exclusions ensures that the proposed exclusion process can be operationalised and provide stability over time. The ECC Model therefore considers all available patient clinical information, including the PDx when determining DCLs and an overall ECCS.

Comparative results have clearly demonstrated that the replacement of PCCL by ECCS within the AR-DRG classification leads to a significantly better performing classification system for acute admitted episodes.

Comparisons show ECCS to perform significantly better than PCCL in explaining cost variation within almost all ADRGs. Although the AR-DRG classification shows slightly better performance overall with fewer ADRG splits (compared to the 5-category ECCS-based splits chosen for use in comparisons), it is clear that the higher performance of the AR-DRG classification is driven by the use of LOS as an ADRG splitting variable within the classification (e.g. same-day DRG splits). This observation, together with the ECCS models' significantly higher performance compared to the PCCL model, demonstrate that if PCCL



were replaced by ECCS, the resulting classification (i.e. AR-DRG V8.0) would significantly outperform AR-DRG V7.0.

The ECCS models' performance on paediatric and geriatric episodes was evaluated and compared to the current PCCL measure and the AR-DRG classification. It was found that ECCS performed well on both cohorts, with exceptional results on paediatric episodes.

For paediatrics the ECCS models demonstrated significantly better performance in terms of minimising bias of cost estimation, showing an adjustment of 0.7 per cent would be required to calibrate the model to paediatric episodes, compared to 3.4 per cent for the AR-DRG classification and 6.6 per cent for the PCCL model.

When restricted to either paediatric or geriatric episodes, the ECCS models show a significantly higher proportion of ADRG-level cost ratios between 0.9 and 1.1, compared to the PCCL model and the AR-DRG classification, demonstrating that ECCS performs consistently well in minimising bias of cost estimation across ADRGs on both cohorts.

The ECCS models do not perform as well on geriatric episodes when compared to performance on paediatric episodes. However, the resulting geriatric cost ratio of 1.015 (or 1.5 per cent increase) is nonetheless a relatively small level of bias of cost estimation, particularly compared to that of the PCCL model and the AR-DRG classification on paediatric episodes.

As part of the project the ACCD was also required to consider the potential role of the COF within the classification. This evaluation process established the difficulty associated with defining what a condition arising during the episode of care meant in terms of its preventability; and how these conditions could be differentiated from other conditions also arising during the acute admitted episode. It was therefore determined that removing ICD codes associated with these conditions from the complexity algorithm reduces the capacity of the classification to explain true cost differences between DRGs.

Changes in clinical care and improvements in data quality over time were identified as necessitating the ongoing evaluation and review of the ECC Model to ensure it is best suited to its proposed role in the AR-DRG classification. The development of a framework that



supports a continuous and systematic approach to evaluation and refinement which encompasses the broader AR-DRG refinement process will be implemented.

In conclusion, the ECC Model is a conceptually based, formally derived and data driven episode clinical complexity measure that has been shown to perform significantly better than the PCCL model within AR-DRG Version 7.0. The performance of the ECC Model within the AR-DRG classification will be the subject of Phase Two of the AR-DRG Review.



11 References

Averill RF, Goldfield N, Hughes JS, Bonazelli J, McCullough EC, Mullin R, Tang AM, Muldoon J, Turner L, Gay J, Neff J, and Sedman A. (2008). 3M™ APR DRG Classification System Version 26.1, *Methodology Overview*. Wallingford, CT: 3M Health Information Systems.

Bottle A, Aylin P. (2011). Comorbidity scores for administrative data benefited from adaptation to local coding and diagnostic practices. *J Clin Epi* 64: 1426-1433.

Brewster A C, Karlin BG, Hyde LA, Jacobs CM, Bradbury RC, Chae YM. (1985). MEDISGRPS: a clinically based approach to classifying hospital patients at admission. *Inquiry: a journal of medical care organization, provision and financing*, 22(4): 377

Cameron AC, Windmeijer FAG (1996). R-Squared measures for count data regression models with applications to health-care utilization. *Journal of Business & Economic Statistics* 14: 209-220.

Cameron AC, Windmeijer FAG (1997). An R-squared measure of goodness of fit for some common nonlinear regression models. *Journal of Econometrics* 77: 329-342.

Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *Journal of Chronic Diseases* 40 (5):373-383.

CMS (Centres for Medicare and Medicaid Services) (2008). [Hospital-Acquired Conditions \(Present on Admission Indicator\)](#), accessed 19 May 2014.

<http://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/HospitalAcq%20Cond/Hospital-Acquired_Conditions.html>.

Commonwealth Department of Health and Aged Care (2000). Development of the Australian Refined Diagnosis Related Groups (AR-DRG) Classification Version 4: Volume 3 Complications and Comorbidities in AR-DRG Classification Version 4.0: Commonwealth of Australia.

Commonwealth Department of Health and Ageing (2008). Australian Refined Diagnosis Related Groups Version 6.0 Definitions Manual: Volume One: Commonwealth of Australia.



Elixhauser A, Steiner C, Harris DR, COFFEY RM. (1998) Comorbidity measures for use with administrative data. *Medical Care* 36(1) 8-27.

Fetter RB, Brand, DA, Gamache, D. (1991) [DRGs: their design and development](http://www.ncbi.nlm.nih.gov/nlmcatalog/9007288) Ann Arbor, Mich: Health Administration Press. <<http://www.ncbi.nlm.nih.gov/nlmcatalog/9007288>>.

Geissler A, Scheller-Kreinsen D, Quentin W, Busse R. (2011). Germany: Understanding G-DRGs. Chapter 14 in Busse, Geissler, Quentin & Wiley, *Diagnosis Related Groups in Europe: Moving towards transparency, efficiency and quality in hospitals*. UK: European Observatory on Health Systems and Policies/ McGraw-Hill.

Gonnella JS, Hornbrook MC, Louis DZ. (1984). Staging of disease: A case-mix measurement. *Journal of the American Medical Association* 251(5): 637-644.

Horn SD, Horn RA, Sharkey PD. (1983). The Severity of Illness Index as a severity adjustment to diagnosis-related groups. *Health Care Financing Review* Suppl: 33-45.

Horn, SD, Sharkey, PD, Chambers, AF, & Horn, RA. (1985). Severity of illness within DRGs: impact on prospective payment. *American journal of public health*, 75(10), 1195-1199.

Jackson TJ, Nghiem HS, Rowell DS, Jorm C, Wakefield J. (2011). Marginal costs of hospital acquired conditions: information for priority setting for patient safety programs and research. *Journal of Health Services Research and Policy* 16(3):141–146.

Jencks SF, Dobson A. (1987). Refining case-mix adjustment: The research evidence. *The New England Journal of Medicine* 317(11): 679-686.

Joint Working Party of the (Australian) Commission (on Safety and Quality in Health Care) and the Independent Hospital Pricing Authority (2013). [A Literature Review on Integrating Quality and Safety into Hospital Pricing Systems and Supplementary Briefing](http://www.safetyandquality.gov.au/national-priorities/jwp-acsqhc-ihpa/). Accessed 19 May 2014. <<http://www.safetyandquality.gov.au/national-priorities/jwp-acsqhc-ihpa/>>.

Kalman I, McCarthy B (2007). *Benchmark studies of DRG systems used in the EU: Experience in Ireland and Germany*. European Commission Twinning Project Report.

Kobel C, Pfeiffer K-P. (2011). Austria: Inpatient care and the LKF framework. Chapter 11 in Busse, Geissler, Quentin & Wiley, *Diagnosis Related Groups in Europe: Moving towards*



transparency, efficiency and quality in hospitals. UK: European Observatory on Health Systems and Policies/ McGraw-Hill.

McNair PD, Borovnicar D, Jackson TJ, Gillett S. (2009a) Prospective payment to encourage system wide quality improvement. *Medical Care* 47(3):272-278.

McNair PD, Luft HS, Bindman AB. (2009b). Medicare's policy not to pay for treating hospital-acquired conditions: the impact. *Health Affairs* 28(5):1485-1493.

National Casemix and Classification Centre (NCCC). (2013). *Australian Refined Diagnosis Related Groups Version 7.0 Definitions Manual: Volume Three (Appendices)*. Wollongong, NSW: NCCC, Australian Health Services Research Institute, University of Wollongong.

National Casemix and Classification Centre (NCCC). (2012a). *The International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, Australian Modification, Australian Coding Standards (8th ed.)*. Wollongong, NSW: NCCC, Australian Health Services Research Institute, University of Wollongong.

National Casemix and Classification Centre (NCCC). (2012b). *The International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, Australian Modification (ICD-10-AM) – Tabular List of Diseases*. Wollongong, NSW: NCCC, Australian Health Services Research Institute, University of Wollongong.

National Centre for Classification In Health (2003). *The Good Clinical Documentation Guide*. Sydney NSW: NCCH, The University of Sydney: Commonwealth of Australia.

Patris A, Gomez S, Mendelsohn M. (2008). Refined French DRG with four severity levels. *AITH Working Paper*, Lyon: Agence Technique de l'Information sur l'Hospitalisation.

Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi JC, Saunders LD, Beck CA, Feasby TA, Ghali WA. (2005). Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Medical Care* 43(11):1130-1139.

Schumacher DN, Parker B, Kofie V, Munns JM. (1987). Severity of illness index and the adverse patient occurrence index: a reliability study and policy implications. *Medical Care* 25(8):695-704.



Smits HL, Fetter RB, McMahon LF. (1983). Variation in resource use within diagnosis-related groups: the severity issue. *Health Care Financing Review*, Suppl: 71-78.



12 Appendices

[Appendix 1:](#) What approaches to DRG development have been taken internationally to account for complications and comorbidities? (Discussion paper)

[Appendix 2:](#) List of Chapter 18 Signs and symptoms codes with DCL in-scope codes identified

[Appendix 3:](#) List of all Chapter 21 codes with Unacceptable PDx codes and DCL in-scope codes identified

[Appendix 4:](#) List of all Unacceptable PDx codes with DCL in-scope codes identified

[Appendix 5:](#) Special case unconditional exclusions: sequelae codes

[Appendix 6:](#) Special case unconditional exclusions: full-time dagger codes

[Appendix 7:](#) Special case conditional exclusions: asterisk codes with single dagger codes

[Appendix 8:](#) List of in-scope diagnosis codes with their Coherent Diagnosis Class



13 List of Figures

Figure 1: The AR-DRG Classification System Development and Refinement Services Governance Structure.....	26
Figure 2: Illustration of ADRG cost profiles – mean, median and interquartile range costs of $E_i(B66)$ for $i = 1, \dots, 10$	32
Figure 3: Illustration of mean diagnosis profile costs of diagnosis x in B66.	33
Figure 4: Scatter plot of diagnosis profile mean cost by CCL value for diseases occurring in $E_2(F14)$	34
Figure 5: Scatter plot of diagnosis profile mean cost by CCL value for diseases occurring in $E_3(F14)$	35
Figure 6: Distribution of Pearson correlation coefficients across ADRGs resulting from comparison of CCLs and diagnosis profile mean costs of episodes with two diagnoses.	36
Figure 7: Illustration of the change in cost associated with x in $E_8(B66)$	38
Figure 8: Illustration of $C_i(A)$ model form with $a = \$1,000$, $b = 1.3$ and three choices of r : $r = 0.8$; $r = 0.88$; and $r = 0.95$	43
Figure 9: The decayed multiplicative model $C_i(G03)$ together with the geometric mean cost of each ADRG profile $E_i(G03)$	45
Figure 10: Distribution of $C_i(A)$ relative change parameter (i.e. b)	47
Figure 11: Distribution of $C_i(A)$ decay parameter (i.e. r).	47
Figure 12: Scatter showing joint distribution of $C_i(A)$ relative change and decay parameters (i.e. b and r).	48
Figure 13: Illustration of the DCL array.....	50
Figure 14: Illustration of aggregation hierarchy for DCL calculation.....	52



Figure 15: Comparison of CCL and DCL correlations with cost among episodes with exactly two diagnoses.	61
Figure 16: Extract of AR-DRG V7.0 manual specifying PCCL formulation.....	64
Figure 17: Episode costs by diagnosis count - E62 ADRG.	65
Figure 18: Episode costs by ECCS - E62 ADRG	66
Figure 19: DRG classification process.	69
Figure 20: Surgical DRG - PDx with a cost at i=1 of more than 3 times greater than the average diagnosis within ADRG F14.	75
Figure 21: Surgical DRG - PDx with a cost at i=1 twice the average diagnosis within ADRG I20.	75
Figure 22: Medical ADRG - where a specific PDx adds additional resources.	76
Figure 23: Other Medical DRG - contains the remainder of PDxs that are not assigned to a specific medical ADRG within the MDC.	76
Figure 24: Comparative difference of RID statistics across ADRGs – ECCS models v PCCL model.	96
Figure 25: Comparative difference of R^2 statistics across ADRGs – ECCS models v PCCL model.	97
Figure 26: Comparative difference of RID statistics across ADRGs – ECCS models v DRG classification.....	98
Figure 27: Comparative difference of RID statistics across LOS-split ADRGs – ECCS models v DRG classification.....	99
Figure 28: Comparative difference of R^2 statistics across ADRGs – ECCS models v DRG classification.....	100
Figure 29: Comparison of ECCS, PCCL and AR-DRG by cost - B02 Cranial Procedures.....	101
Figure 30: Comparison of ECCS, PCCL and AR-DRG by cost - G03 Stomach, Oesophageal and Duodenal Procedures.	102



Figure 31: Comparison of ECCS, PCCL and AR-DRG by cost - H01 Pancreas, Liver and Shunt Procedures.	103
Figure 32: Comparison of ECCS, PCCL and AR-DRG by cost - T60 Septicaemia.	104
Figure 33: Distribution ADRG-level cost ratios across each model on paediatric episodes..	107
Figure 34: Distribution ADRG-level cost ratios across each model on geriatric episodes.....	110



14 List of Tables

Table 1: Summary of final data used in the project (private and public establishments).	29
Table 2: The decayed multiplicative model $C_i(G03)$ together with the geometric mean cost of each ADRG profile $E_i(G03)$	46
Table 3: Aggregation hierarchy for a diagnosis-ADRG combination (x;A).....	54
Table 4: Notation used in calculation of cumulative relative cost.	55
Table 5: Definition of the sequence of mutually exclusive E_i episode sets.....	57
Table 6: Summary of aggregation levels at which the $\bar{C}(x;A)$ are calculated.	59
Table 7: Proportion of DCLs by diagnosis group.....	60
Table 8: Percentage breakdown of episodes by rounded ECCS.	67
Table 9: Examples of patient diagnoses within pre-coordinated codes with multiple clinical concepts.	73
Table 10: Characteristics of patients in five scenarios.	89
Table 11: Comparative summary of the performance of ADRG splitting models.....	94
Table 12: Summary of percentage distribution of episodes by PCCL, DRG and ECCS categories.....	95
Table 13: Breakdown of ADRGs within each model by number of model categories.	95
Table 14: Comparative summary of the performance of ADRG splitting models using R2 and RID - paediatric episodes.	106
Table 15: Comparison of cost ratios on paediatric episodes.....	107
Table 16: Comparative summary of the performance of ADRG splitting models using R2 and RID - geriatric episodes.	108
Table 17: Comparison of cost ratios on geriatric episodes.	109



Table 18: Variation of DCLs - 2009-10 to 2011-12 full sample vs. 50 x 2009-10 to 2011-12 random 95% subsamples.	115
Table 19: Variation of DCLs - 2008-09 to 2010-11 full sample vs. 50 x 2008-09 to 2010-11 random 95% subsamples.	116
Table 20: Variation of DCLs - 2009-10 to 2011-12 full sample vs. 2008-09 to 2010-11 full sample.	117
Table 21: Illustration of the method changing DCLs based on halving the difference between original and revised calculation.	118
Table 22: Illustration of resulting changes to DCLs if a half-difference method is used to update the DCL matrix.	119