B·H·F

Board of Healthcare Funders

ALTERNATIVE

Reimbursement Models

SPECIAL REPORT

Developing Alternative Reimbursement Models for Hospital Admissions related to COVID-19

COLLABORATIVE WORK

of the Industry Technical Advisory Committee (ITAC)

EXECUTIVE SUMMARY

The BHF was tasked with developing an alternative reimbursement model (ARM) for COVID-19 hospital admissions to assist member schemes with the management of the risk associated with COVID-19 claims – this report focuses exclusively on hospital costs. The BHF invited the industry to participate in a task group for this purpose.

A literature review was undertaken to identify methodologies that had been applied to this problem statement in the past. The findings were used to inform the methodology applied.

Data were requested and received from schemes that account for approximately 37% of the industry. Some modifications were necessary to ensure that the data were reflective of the claims being considered.

Following some initial modelling it was noted that a global fee structure was unlikely to yield results, given the volatility in total hospital admission costs; so the approach was changed to focus on the creation of *per diem* rates (fixed rates per day).

Initially it was thought that these could be created based on ward type; however, as a result of how the data were collected, it was only possible to obtain consistent results for general ward admissions. This was in part due to the mix of costs if someone was in general ward, high care and ICU over the course of their treatment.

Based on the modelling, it appears possible to establish *per diem* rates for COVID-19 hospital admissions. The modelled *per diem* rates quite closely match the observed daily rates for COVID-19 hospital admissions.

Given the modelling outcomes, the daily rates would be varied by level of care, hospital group and the presence of a relevant pre-existing condition.

There are a number of limitations to the process and learnings for future exercises, such as improving the data specification to allow length of stay to be determined more accurately by level of care.

There is an online tool showing the model results that has been prepared to allow readers to view the model results. This is available on the following link: https://www.bhfportal.co.za/bhf_global/HOSP/

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ITAC Team Members

Table 1: ITAC team members

Name	Organisation
Paresh Prema	Independent
Adam Lowe	NMG Consultants and Actuaries
Aphiwe Baleni	Percept Actuaries and Consultants
Charlton Murove	Board of Healthcare Funders
Dr Rajesh Patel	Board of Healthcare Funders
Lesley Mogano	Board of Healthcare Funders
Craig Getz	Insight Actuaries and Consultants
Evan Bradley	3ONE Consulting Actuaries
Dr Jenni Noble-Luckhoff	Medscheme
Linda Webb	Alexander Forbes
Martin Coxon	Percept Actuaries and Consultants
Sarah Bennett	Medscheme
Shivani Ranchod	Percept Actuaries and Consultants

Acknowledgements

This project commenced as a response to the ongoing impact of the COVID-19 pandemic in South Africa and in particular its management from a health financing perspective.

It is for this reason we would like to thank the team at the BHF for their initiative in developing an ARM for COVID-19. Their support from an infrastructure and facilitation point of view was invaluable in allowing the team to develop this model. Furthermore, the model would not be possible without the technical input of the team from the BHF namely, Mr Charlton Murove and Dr Rajesh Patel.

The ITAC was created to undertake work that would provide rigorous and effective solutions to challenges in the healthcare landscape. We would like to thank the members of the team (listed in the table above) for their valuable contribution to conceptualising, developing and finalising this model. We would like to extend special thanks to Mr Paresh Prema who is the chair of the ITAC and who coordinated the efforts of the team.

We would also like to extend our gratitude to Prof. Eustasius Musenge who peer reviewed the modelling approach and this report.

This model would not be possible without medical schemes data and we would like to thank the medical schemes and administrators that contributed resources and data for its development.

1. Introduction

When the COVID-19 pandemic started, the possible exposure of medical schemes to the costs was identified as a risk. In addition to this, COVID-19 was classified as a prescribed minimum benefit (PMB), which increased this risk. As a result, the idea of an ARM for COVID-19 hospital admissions was considered to be a possible way of mitigating the risk. ARMs include a range of alternative structures to the current fee-for-service model primarily used. Types of ARMs include:

- Per day (*or per diem*) rates: A fee per day in hospital
- Fixed fees or global fees: This is where a fee is agreed upon that encompasses all costs related to a specific treatment or health event
- Pay for performance or value-based reimbursement: These models measure outcomes and create incentive structure for positive outcomes.

Initially a global fee was considered to be the best approach. It was decided that this should be researched and developed by the BHF, which would facilitate a task team for this purpose.

During October 2020 a terms of reference was approved and the task team, which consisted of actuaries and two clinicians from across the industry, assembled. The task team put together a specification for the data required to develop the model and this went out to all schemes who are members of the BHF. More detail regarding this data specification is included in Annexure 2.

The task team, when debating the structure of the model, agreed on two approaches to the creation of an ARM/global fee, namely:

- *Top-down* – this team considered the data and looked for a way to create homogenous groups where there was a consistent claim amount that could form the base of the global fee
- *Bottom-up* – this team considered COVID-19 and the treatment protocols that could be built up from first principles to create a treatment basket and resulting global fee.

A concern was raised about role of the BHF in assembling a technical team given that the BHF represents member schemes. The concern was that the BHF is pooling technical skills across the funding industry, and that this could be perceived as schemes combining resources and working on a project that could influence scheme-provider negotiations. This could be seen as anticompetitive.

It is important for ITAC to address this concern directly in terms of intention, independence and impact. Firstly, from the perspective of intention: ITAC was created by the BHF to serve the wider healthcare industry. BHF itself employs industry experts, including a medical doctor, with many years of experience in the funding industry, and a qualified actuary, who contributed their knowledge and skill to this report. The Competition Act does not prohibit research into fee structures or other technical issues affecting medical schemes or the health sector. Not every collaborative effort between competitors is anticompetitive. Pooling knowledge and expertise to add to the body of knowledge within an industry for use by individual players as they see fit is not anticompetitive. The people who work in the industry are the very experts whose knowledge and experience add credibility to such research. The BHF is not only an industry representative association. It also serves as a repository for industry-level knowledge and expertise that is valued by its members.

The Competition Act prohibits restrictive horizontal practices. Not all horizontal practices are restrictive. The Act defines what is meant by restrictive horizontal practices in section 4. Basically these are practices that *substantially* prevent or lessen competition. This paper does not do so. It is a research report created by a group of experts and as such is not of a commercial nature. It is not directly or indirectly fixing any fees or trading conditions or engaging in concerted practices that have the effect of substantially lessening or preventing competition in a market. Its goal is to present information that is beneficial to the industry as a whole.

Secondly, ITAC members act in their professional capacity. The majority of the members in the task team do not represent particular medical schemes but are volunteering their technical competencies. The role played by the BHF was one of initiating the project and facilitating the work done for the benefit of the wider industry. Many of the members of ITAC are actuarial consultants with both funder and provider clients.

In terms of impact, the work done will be shared publicly and it is envisioned that it can empower and be used by different healthcare stakeholders to create alternative reimbursement arrangements that can improve the resilience of our healthcare system in the fight against COVID-19. The work done will also provide a useful framework that the industry can build on when creating ARMs for other conditions.

It is also useful to distinguish between reimbursement structures and the prices associated with those structures. The focus of ITAC has been to consider questions relating to reimbursement structures. It is in the interests of both funders and providers to have coherent reimbursement structures that reflect underlying clinical dynamics. Reimbursement fragmentation creates system complexity and cost, and undermines the likelihood of funders and providers aligning incentives.

2. Literature review

2.1. COVID-19 related papers

Given the nature of the COVID-19 pandemic, much research focus has been placed on understanding various aspects of the disease and its treatment. However, since managing the spread and severity of COVID-19 infections is a public health issue in most countries, the focus has been on the utilisation of resources rather than cost per case (as is often analysed in private sector work). Thus most of the research relevant to this initiative focuses on the factors causing longer lengths of stay, higher levels of care, higher mortality rates or the greatest utilisation of other hospital resources.

As private sector utilisation of these resources incurs a charge, it is reasonable to expect that the factors impacting resource utilisation will also impact the cost per admission. The most tenuous link is likely that to mortality (which is used as the dependent variable in a number of COVID-19 hospitalisation studies), but it could be argued that patients who eventually die are among the sickest and most costly patients to be treated.

A number of studies have been published analysing the impact of various factors on the severity of infection and the consequent resource utilisation, including the following:

- A study from London (1) identified age, gender (specifically a higher risk for males) and other clinical presentation factors as the key factors associated with 30-day mortality for hospitalised patients;
- A Chinese study (2) used a logistic regression model for 'severe symptoms' to identify an increasing probability of severe symptoms (and hence higher resource utilisation) associated with increasing age and male gender;
- A study from Norway (3) modelled probabilities of hospitalisation, the development of severe disease (modelled as ICU admission) and death by various risk factors, of which age was the most significant factor across the variables, while male gender and number and severity of comorbidities (notably diabetes and heart failure) were also important risk factors;
- A second Chinese study (4) used ICU admission as a proxy for severity of disease and found the key risk factors to be age, the presence of comorbidities (notably hypertension, cerebrovascular disease and diabetes) and some specific clinical markers;
- A Danish study (5) found that age, comorbidity index and specific clinical markers were the most significant predictors of mortality in patients admitted with COVID-19; and
- Another study from China (6) using mortality as its dependent variable, identified age, comorbidities (most notably diabetes, hypertension and chronic cardiac disease) and a series of key clinical markers as the most significant predictors of death in severe cases of COVID-19 infection.

One piece of South African research (7) attempted to consider average costs per day and lengths of stay for COVID-19 patients. Through an analysis of published research and a number of assumptions, the authors used the below set of 'model variables' in their study, which aimed to assess the cost-effectiveness of using ICU care relative to general ward care for public sector COVID-19 patients:

1. Mortality rates
2. Utilisation of inpatient days
3. Disability weights associated with severity of disease
4. Unit cost per general ward day and per ICU day

This was also split by private and public sectors.

The National Department of Health (NDoH) of South Africa released guidelines on public private partnerships in response to COVID-19. The guidelines included modelled reference *per diem* rates for admitted patients in private hospitals (8). These *per diem* rates were determined for a weighted average of high care and ICU beds.

2.2. Other papers relating to modelling of hospital costs

In general, it is acknowledged that because of the skewness of healthcare data, ordinary least squares (OLS) regression modelling is unsuitable for such data. As such a number of alternative methods are suggested to model healthcare data, including OLS models on transformed data (commonly a log transform), generalised linear models (GLMs), other parametric models outside of the GLM family, and a sequence of more complex non-parametric approaches. Notably:

- A recent paper produced through the European Union's Horizon 2020 research and innovation programme (9) noted that transformed OLS models required complex retransformation in many cases; GLM models work effectively on healthcare data as long as the distribution is specified correctly; a few alternative approaches were noted and a new semi-parametric GLM-based approach proposed;
- An older paper summarising statistical methods for analysing healthcare costs (10) noted that simpler methods, i.e. those based on OLS, either raw or transformed data, work well with large samples, while the GLM approaches again work well when the distribution is correctly specified; the more complex approaches require significant time and expertise to be applied correctly;
- An Italian paper analysing healthcare cost modelling (11) noted that where no zero costs and no censoring of the data existed, the log-gamma GLM was the most favoured approach; other approaches are suggested where zero-cost data points exists, most notably multi-step modelling; and
- Another study using population healthcare costs as its base and intending to deal with the analysis of skewed data (12) concluded that in most cases a gamma-based GLM model performed well, with OLS-based models improving with larger sample sizes, while the other models tested didn't perform as well as the gamma GLM.

2.3. Trimming data

When analysing hospital cost-related data, it is common practice to remove outliers in terms of cost and length of stay.

The L3H3 method of trimming length of stay is commonly used to determine trim points (12-14). Tukey's rule has been extensively used to determine outliers in many contexts including healthcare (16–19).

3. Data

The data received came from six schemes that cover approximately 37% of medical scheme lives. Data were collected for hospitalised beneficiaries from 1 January 2019 to 31 December 2020. The data provided included ICD10 codes. COVID-19 claims were identified as claims with either the primary or the secondary diagnosis being classified as U07.

A summary of the data received in respect of COVID-19 admissions is detailed below.

Table 2: COVID-19 admissions

Gender	Number of admissions	Average length of stay	Average cost per admission	Standard deviation of cost per admission	Total paid R'million
Female	13 845	8.79	R90 472	208 620	1 253
Male	9 573	9.50	R114 176	258 024	1 093
Total	23 418	9.08	R100 162	230 427	2 346

The data included all admissions to the data extraction date. The inherent challenge is that some claims relating to such admissions may not have been submitted and in instances where claims have been submitted, they may not have been paid. The section below details adjustments made to the data to partly account for this.

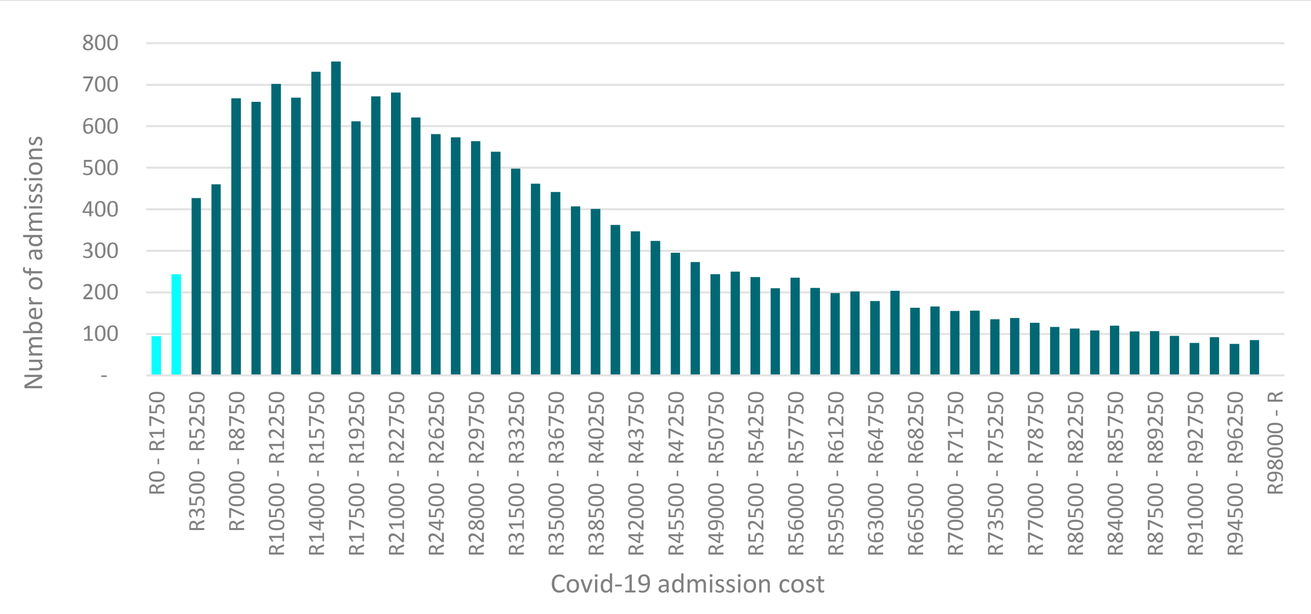
Additional claims data were requested relating to asthma, bronchitis, COPD, pneumonia and respiratory failure. The intention was that if COVID-19 cases were included in these data they would be identifiable and could be recategorised. Following a review of the data it was determined that this would not be possible based on the level of detail requested. Therefore, only clearly defined COVID-19 claims were included.

3.1. Adjustments to data

Due to the data collection approach, it was known that there would be some incomplete admissions, where claims in respect of the admissions may not yet have been fully received or settled. In order to remove these outliers, claims with a total amount paid of less than R3 500 were excluded. This level was chosen as it approximately reflects the minimum cost of a one-day admission.

The below frequency graph was used to support the exclusion of claims less than R3 500. The graph considers the spread of cost per admission for all COVID-19 admissions with a cost less than R99 750.

Figure 1: COVID-19 hospital cost per admission



The graph above confirms that those admissions below R3 500 appear to be outliers.

The effect of the exclusions was a reduction from 23 418 to 23 080 admissions (i.e. 338 admissions were excluded). The revised summary of the data is shown below.

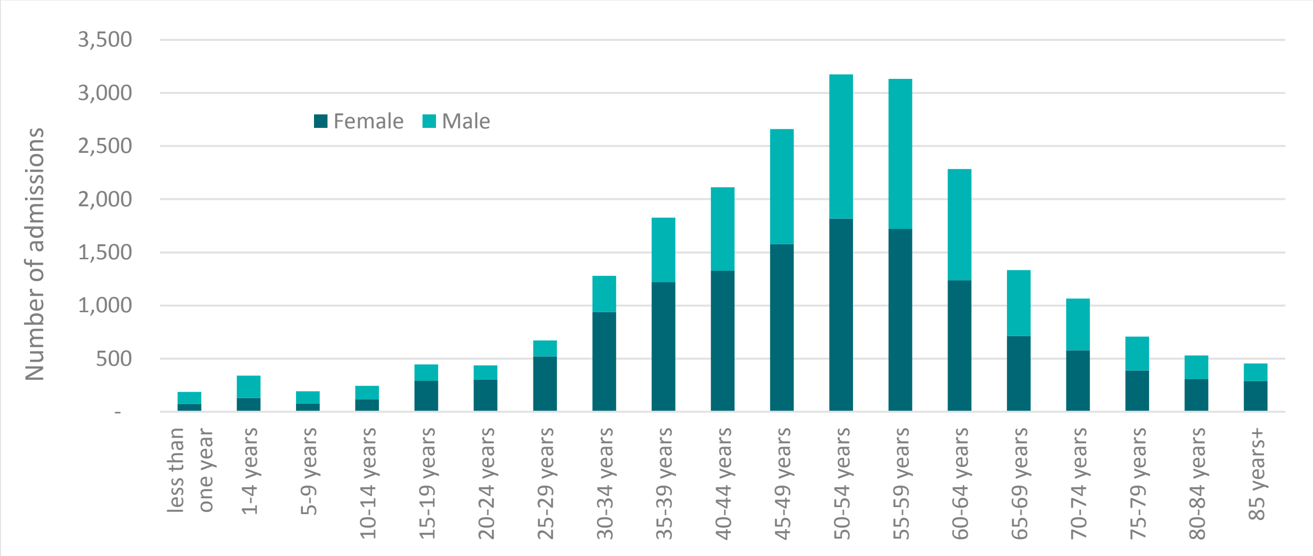
Table 3: Summary of COVID-19 admissions after excluded data

Gender	Number of admissions	Average length of stay	Average cost per admission	Total paid R'million
Female	13 629	8.89	R91 871	1 252
Male	9 451	9.60	R115 622	1 093
Total	23 080	9.18	R101 597	2 345

3.2. Data summaries

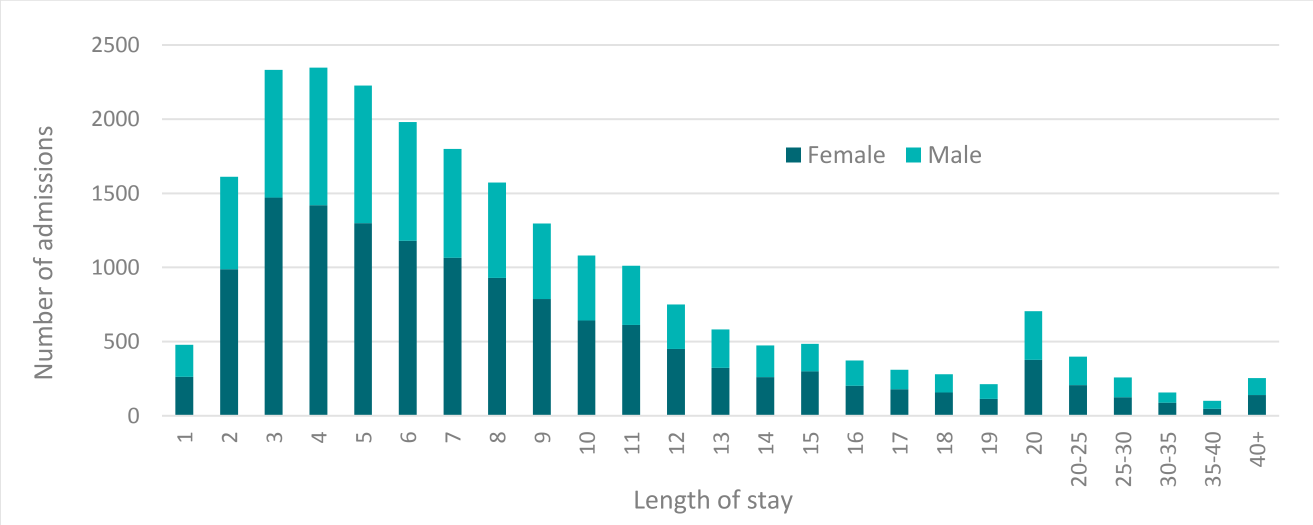
The split of the admissions by age and gender are shown below. Across all age bands, from 15 years and above, there were more females than males admitted.

Figure 2: COVID-19 admissions by age and gender



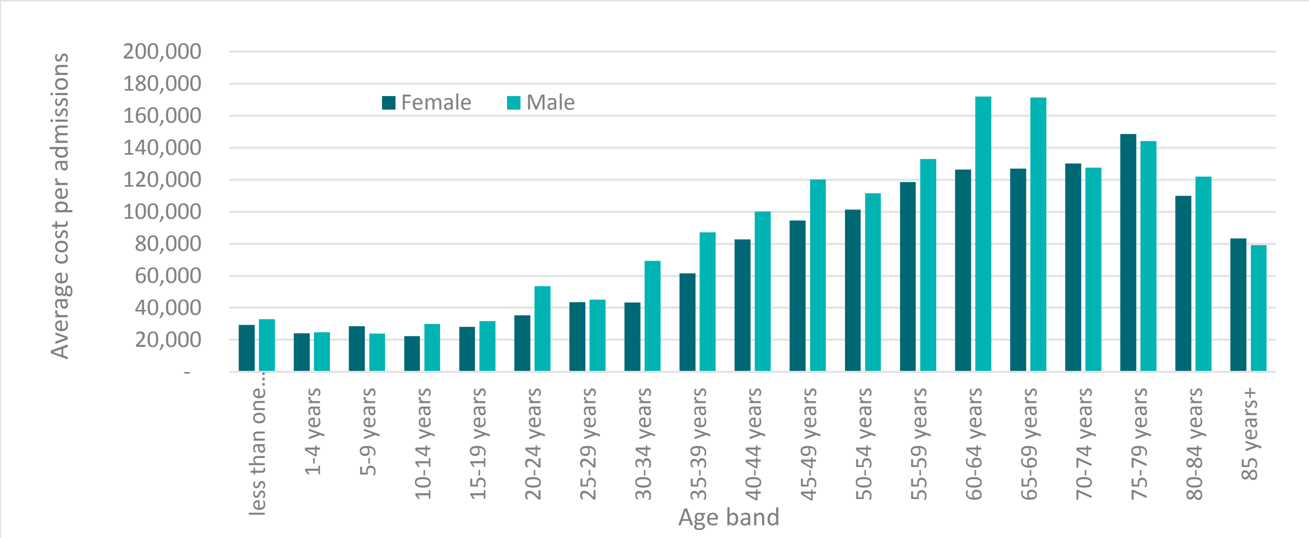
The graph below shows the split of the length of stay by gender. By far the majority of the admissions were for less than five days. The length of stay has a significant tail, with a notable number of admissions having particularly long lengths of stay.

Figure 3: COVID-19 admissions by length of stay and gender



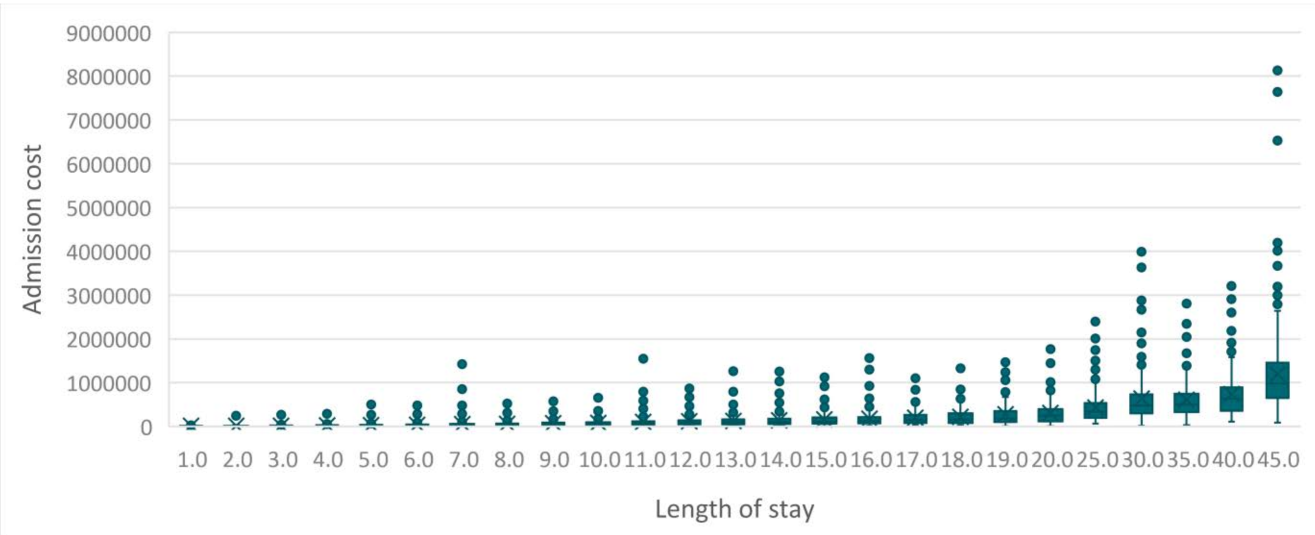
The graph below shows the average cost by age and gender. Across the majority of the age bands the average cost for males was higher than females. This is consistent with the research discussed as part of the literature review, where males were generally seen to exhibit a higher probability of severe outcomes.

Figure 4: COVID-19 average cost per admission by age and gender



Ultimately, to determine the viability of any form of ARM, there needs to be some consistency and predictability in the treatment costs. This uniformity can be achieved through some segmentation of the data but it needs to be identified. The graph below shows that even with length of stay the range of costs is very big. This raised some concerns regarding the viability of the exercise.

Figure 5: Spread of COVID-19 admission cost by length of stay



4. Methodology and assumptions

This chapter summarises the methodology and assumptions employed in the study.

4.1. Identification of chronic beneficiaries

Chronic beneficiaries were identified by the schemes and administrators submitting data utilising criteria published by the Council for Medical Schemes (CMS) to identify PMB chronic conditions.

The latest published guidelines date from 1 January 2018 and are available at:
https://www.medicalschemes.co.za/publications/#2009-2567-wpfd-2020-asr-utilisation-system-current.

4.2. Classification of hospital claims

4.2.1. Hospital claims

Hospitals claims were identified using the discipline code provided by the Practice Code Numbering System (PCNS), as per table below.

Table 4: Facility discipline codes

Discipline Code	Hospital
047	Drug & Alcohol Rehab Facilities
049	Sub-Acute Facilities
055	Mental Health Institutions
056	Provincial Hospitals
057	Private Hospitals ('A' - Status)
058	Private Hospitals ('B' - Status)
059	Private Rehab Hospitals (Acute)
076	Unattached Operating Theatres / Day Clinics
077	Day Clinics
079	Hospices

4.2.2. Level of care

The hospital admissions were classified into three categories for the level of care: general ward, high-care and ICU. The procedure codes were used to classify the level of care. The following process was followed:

- Default level of care was general ward
- Instances where there were codes reflecting high-care were classified as such. This means if an admission started in general ward and transitioned to high-care the entire admission was classified as high-care
- Finally, if there were codes reflecting an ICU admission this was classified as an ICU admission. The ICU admission may include days of general ward and high care.

Table 5: In-hospital level of care procedure codes

Procedure Code	Level of Care
58200	ICU
58201	ICU
58202	ICU
57200	ICU
57201	ICU
57202	ICU
58215	High-care
58216	High-care
58217	High-care
57215	High-care
57216	High-care
57217	High-care

4.3. Trim points

Effective ARMs have trim points included in the model. Trim points are used to eliminate data points that are outliers and have the potential to significantly alter results if included in the model. Trimming removes admissions that are either too short or too long leaving admissions that allow a better statistical fit. Trimming is done on both length of stay and the cost of admissions.

For this analysis Tukey's method was used for trimming data for hospital costs and the L3H3 method was used for the length of stay. The data did not follow any underlying distribution, thus making Tukey's method and L3H3 method appropriate (Tukey's method is ideal for skewed data distribution).

4.3.1. Trimming using hospital costs

High-cost outlier limit is calculated as follows:

Q3 + 1.5*(Q3 – Q1)

Where Q3 is the third quartile of hospital costs and Q1 the first quartile of hospital costs.

If an admission cost more than this limit, it is excluded. The limit was calculated to be R182 220.

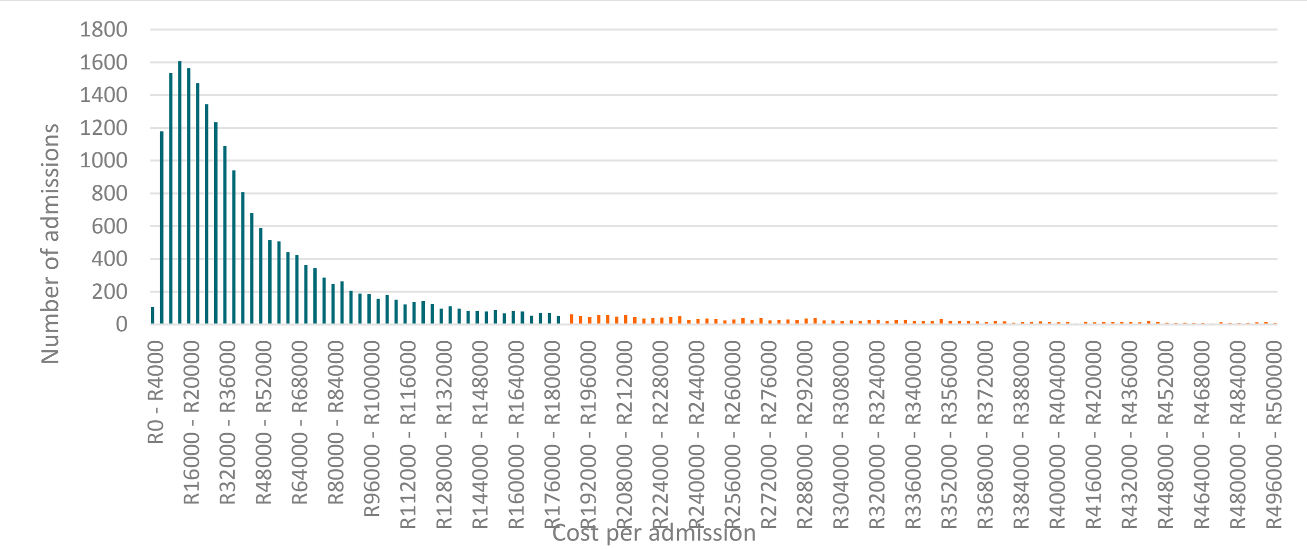
The low-cost outlier limit is calculated as follows:

Q1 - 1.5*(Q3 – Q1)

Similarly, if an admission costs less than this limit, it is excluded. This limit was not applicable, as the calculated value was lower than zero.

The graph below highlights the expensive admissions that were considered outliers and excluded.

Figure 6: COVID-19 admission after trimming by cost of admission



In addition to the bars highlighted in orange that were excluded as outliers, a further 854 admissions exceeding R500 000 were excluded.

4.3.2. Trimming using hospital length of stay

High length of stay outlier limit is calculated as follows:

300% of the Average Length of Stay

If an admission is longer than this limit, it is excluded. The limit was calculated to be 27.5 days.

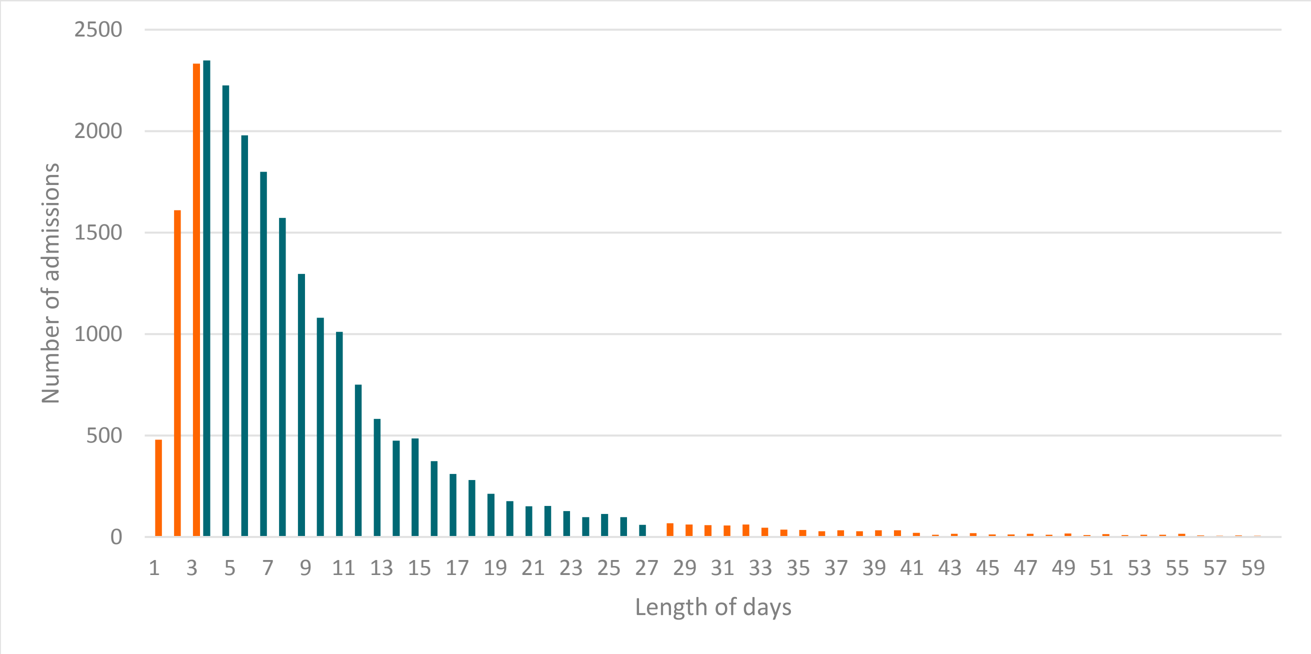
The low length of stay outlier limit is calculated as follows:

1/3 of the Average Length of Stay

Similarly, if an admission is less than this limit, it is excluded. The limit was calculated to be 3.1 days.

The graph below highlights the admissions that were considered outliers and excluded as a result of the length of stay.

Figure 7: COVID-19 admission after trimming by length of stay



In addition to the bars highlighted in orange that were excluded as outliers, a further 86 admissions where the length of stay exceeded 60 days were excluded.

4.4. Summary of data after trimming

A summary of the data received in respect of COVID-19 admissions after trimming is detailed below.

Table 6: COVID-19 admissions after trimming data

Gender	Number of admissions	Average Length of stay	Average cost per admission	Standard deviation of cost per admission	Total paid R'billion
Female	9 418	8,33	R51 471	37 519	0,485
Male	6 324	8,30	R54 444	39 616	0,344
Total	15 742	8,32	R52 665	38 402	0,829

The number of admissions drops to 15 742 from 23 080 while the change in the average length of stay is a drop of 0.86 day. The significant difference is in the average cost per admission, which reduces by almost 50% from R101 597 to R52 665. This does confirm that there is significant variability in the total benefits paid for COVID-19 admissions.

COVID-19 admissions in general ward (no high care or ICU).

Table 7: COVID-19 admissions after trimming data for general ward admissions only

Gender	Number of admissions	Average length of stay	Average cost per admission	Standard Deviation of cost per admission	Total paid R'million
Female	7 695	7,97	R42 402	29 438	326,280
Male	4 926	7,92	R43 291	30 161	213,253
Total	12 621	7,95	R42 749	29 724	539,533

COVID-19 Admissions with stay in high-care ward, i.e. possibly some stay in general ward but no ICU stay.

Table 8: COVID-19 admissions after trimming data for high-care and possibly general ward admissions.

Gender	Number of admissions	Average length of stay	Average cost per admission	Standard deviation of cost per admission	Total paid R'million
Female	1 163	9,90	83 100	40 431	96,646
Male	870	9,76	82 033	40 521	71,369
Total	2 033	9,84	82 643	40 463	168,014

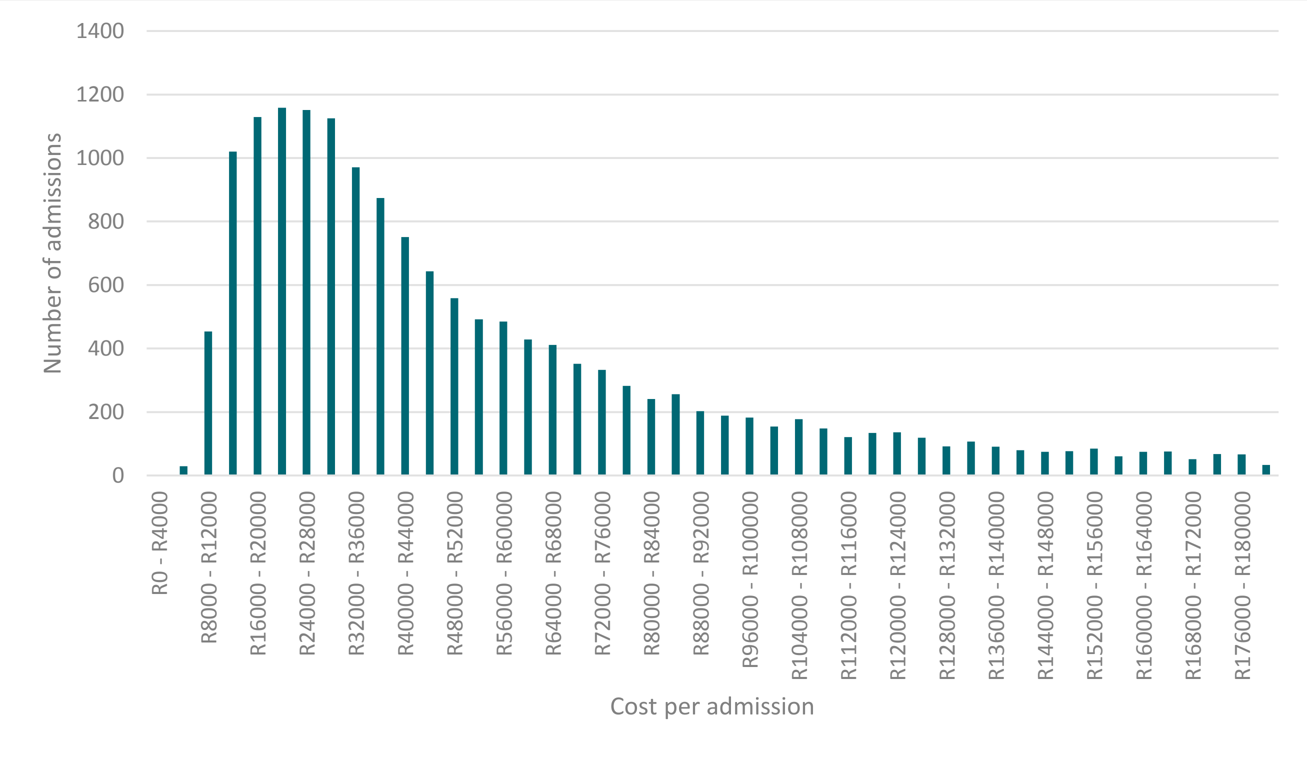
COVID-19 Admissions with stay in ICU, i.e. possibly some stay in general ward and high care.

Table 9: COVID-19 admissions after trimming data for ICU and possibly high-care/general ward admissions

Gender	Number of admissions	Average length of stay	Average cost per admission	Standard deviation of cost per admission	Total paid R'million
Female	560	9,96	110 408	41 167	61,828
Male	528	9,48	113 032	41 993	59,681
Total	1 088	9,73	111 681	41 572	121,509

4.4.1. *Spread of admission costs*

Figure 8: COVID-19 admissions by cost per admission after trimming

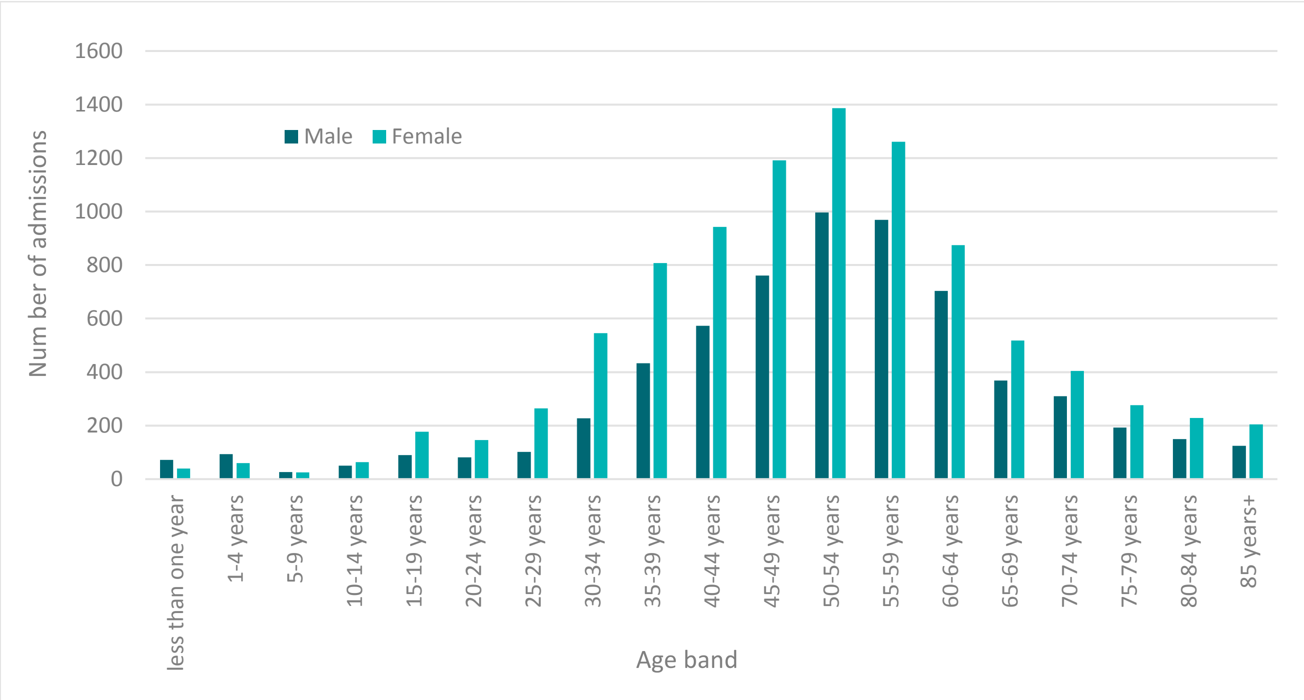


The graph above shows that there was a very wide spread in the cost of COVID-19 admissions, with a large tail of more expensive admissions, even after applying the trimming.

The distribution above appears to indicate that a relatively large volume of admissions occurred at between R10 000 and R30 000. However, the overall average cost per admission was approximately R52 700 due to a relatively large number of expensive admissions. The median cost per admission was approximately R39 800.

4.4.2. Number of admissions by age and gender

Figure 9: COVID-19 admissions by age and gender after trimming data



There were consistently more admissions for females than for males, excluding for young children (below 10 years). There was a total of 9 418 admissions for females relative to 6 324 admissions for males.

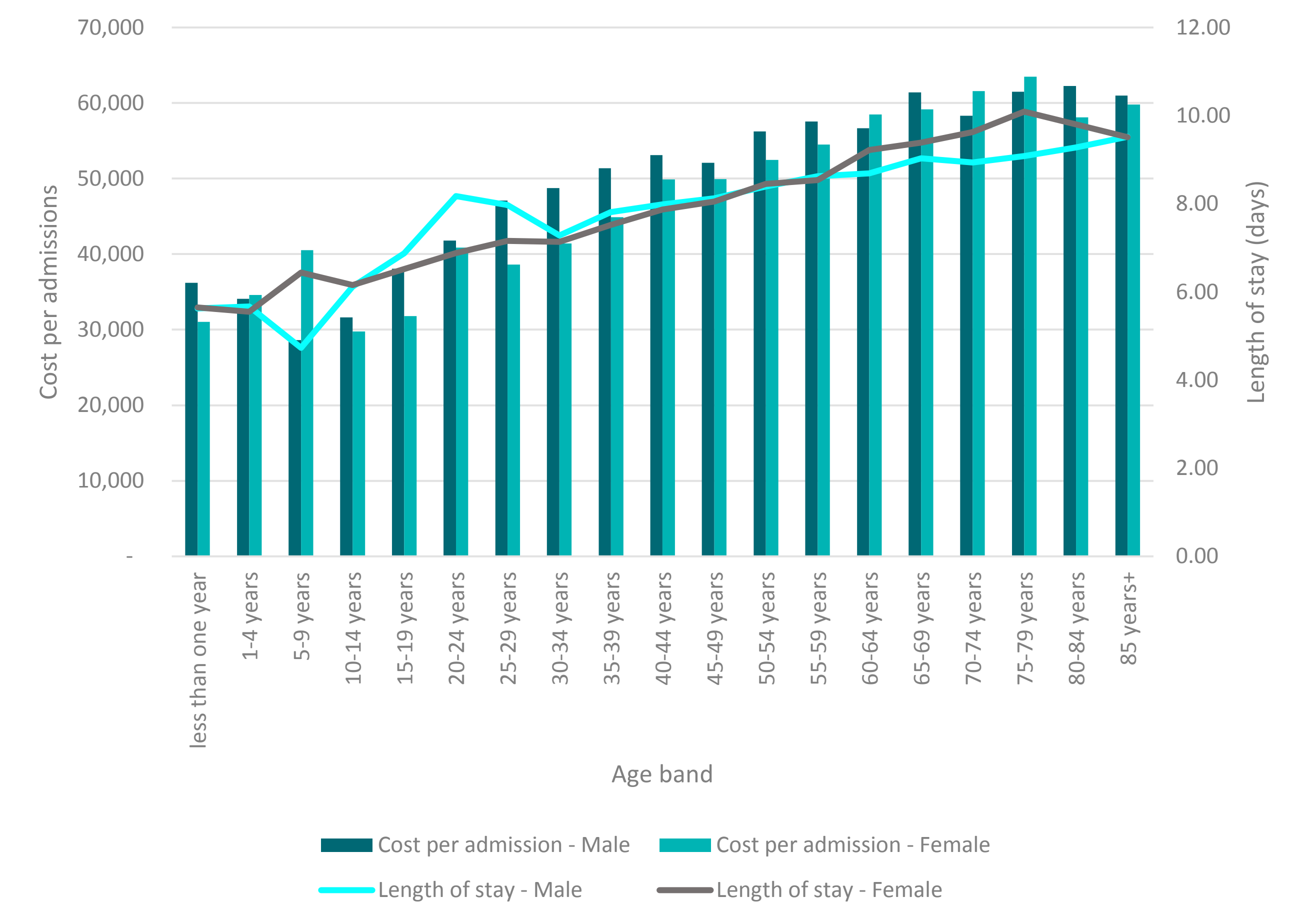
Most admissions were evident in the age band of 50-54 years, with an overall average age for those admitted of 51.0 years. This was similar for both males and females.

The large spread in admission costs, as discussed in the previous section, causes difficulty in structuring an ARM or global fee arrangement, as it indicates that the level of stability or predictability in treatment costs may be low.

Potential drivers of data volatility were, however, analysed to assess whether there was a clear and predictable relationship between costs and these factors, so that segmenting the data by these factors might allow for stable predictions.

4.4.3. Cost experience by age and gender

Figure 10: COVID-19 admissions after trimming data



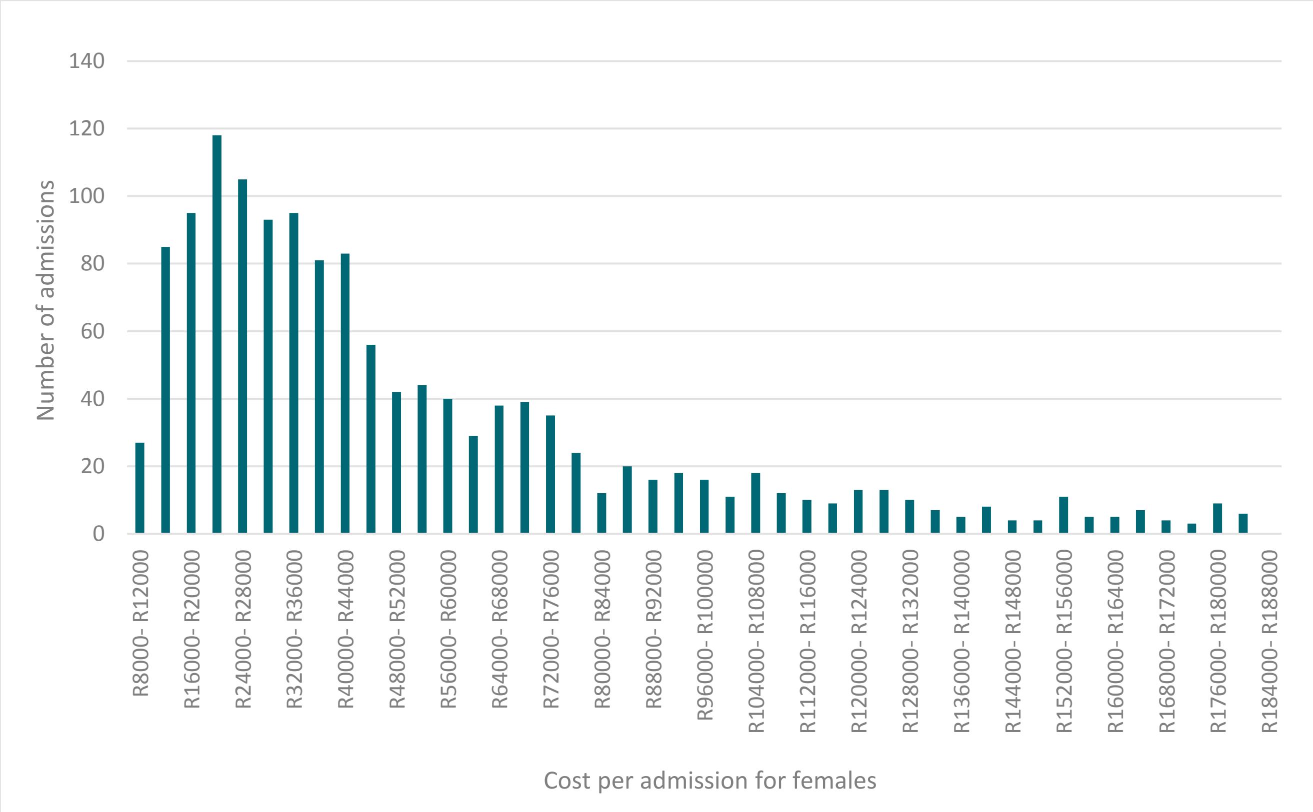
The cost per admission was seen to increase consistently with age, driven primarily by an increase in average length of stay.

The average cost per admission for males was consistently slightly higher than for females, despite a similar average length of stay. This indicates a higher level of severity and cost per day for males, on average.

There appears to be a relatively consistent relationship between age, gender and expected cost, which may improve the predictability of cost per admission. The following graph examines this through considering the spread of admission costs for females aged 50-59 years (i.e., removing the impact on cost volatility driven by age and gender).

4.4.4. *Spread of cost (females aged 50-59)*

Figure 11: COVID-19 spread of cost per admission for females

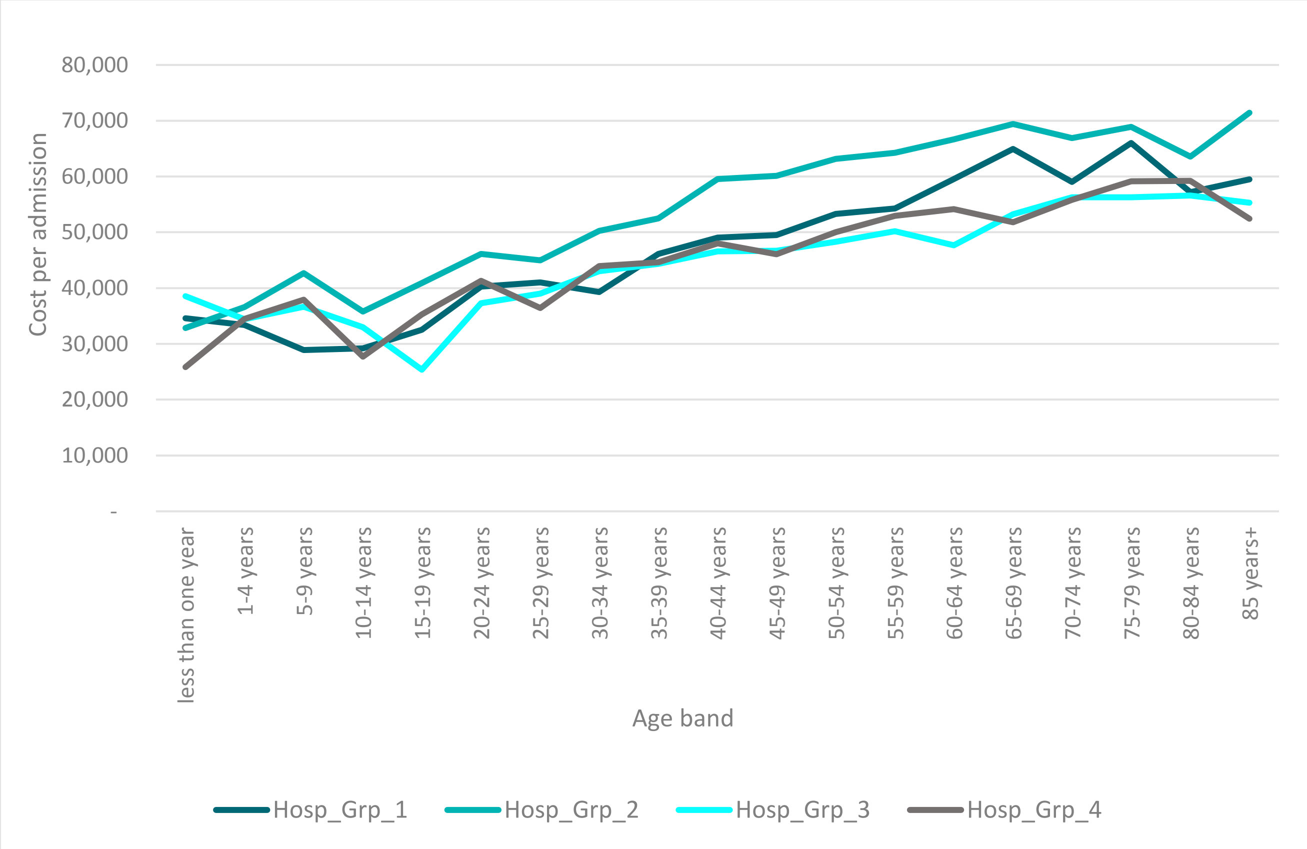


Despite stripping the impact of gender and age from the claims volatility, there remains a large spread of admission costs.

Another area of potential cost volatility relates to varying experiences within hospital groups. The graph below examines the cost experience by hospital group by age band.

4.4.5. Cost experience by hospital group

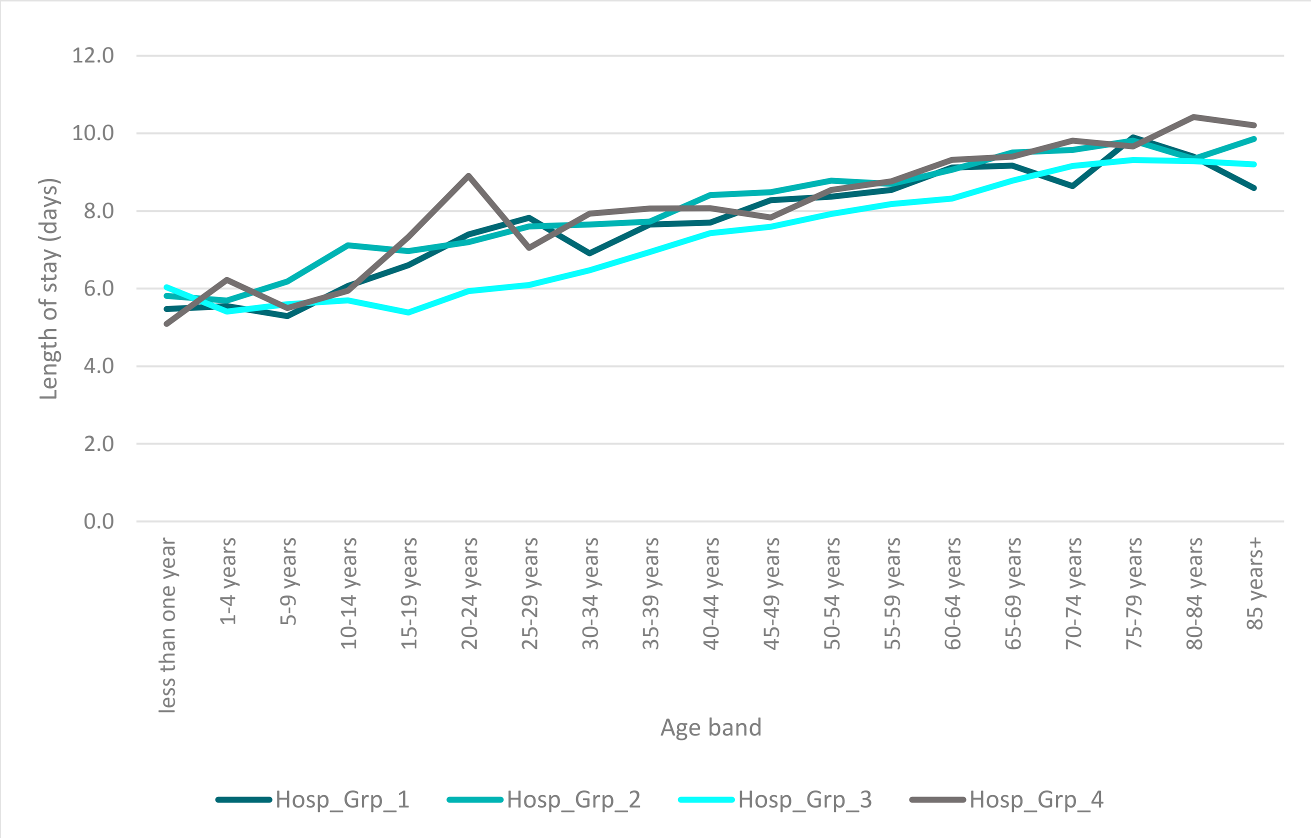
Figure 12: COVID-19 cost per admission by hospital group and gender



The graph above shows that hospital group 2 consistently tends to prove more expensive, while hospital group 1 also proves slightly more expensive than hospital groups 3 and 4 for patients of older ages. However, none of the claims have been adjusted for possible different underlying risk profiles of the members being treated or differences in tariffs of the various schemes. This would have an impact on the actual differential.

4.4.6. Length of stay by hospital group

Figure 13: COVID-19 length of stay per admission by hospital group and gender



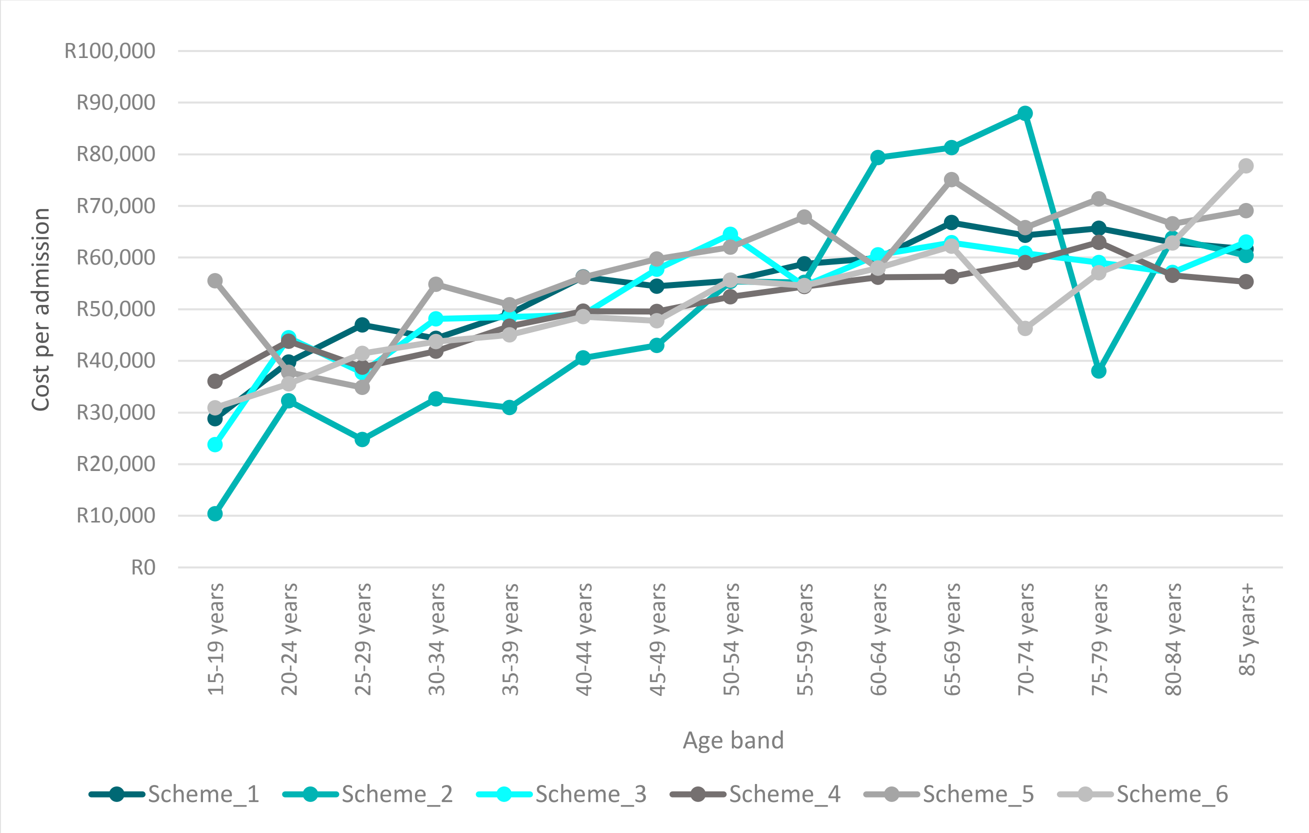
The length of stay by hospital group is relatively consistent. However, hospital group 3 tended to exhibit lower average lengths of stay.

This also demonstrates that the higher average costs in hospital group 2 were not explained by higher lengths of stay, but by a higher cost per day.

These examples of variability by hospital group increase the difficulty of establishing a universal ARM or global fee that would be accepted by all hospital groups, without driving up industry costs in the event that the fee were to be set at the level of costs evident within hospital group 2.

4.4.7. Cost experience by medical scheme

Figure 14: COVID-19 cost per admission by scheme and gender



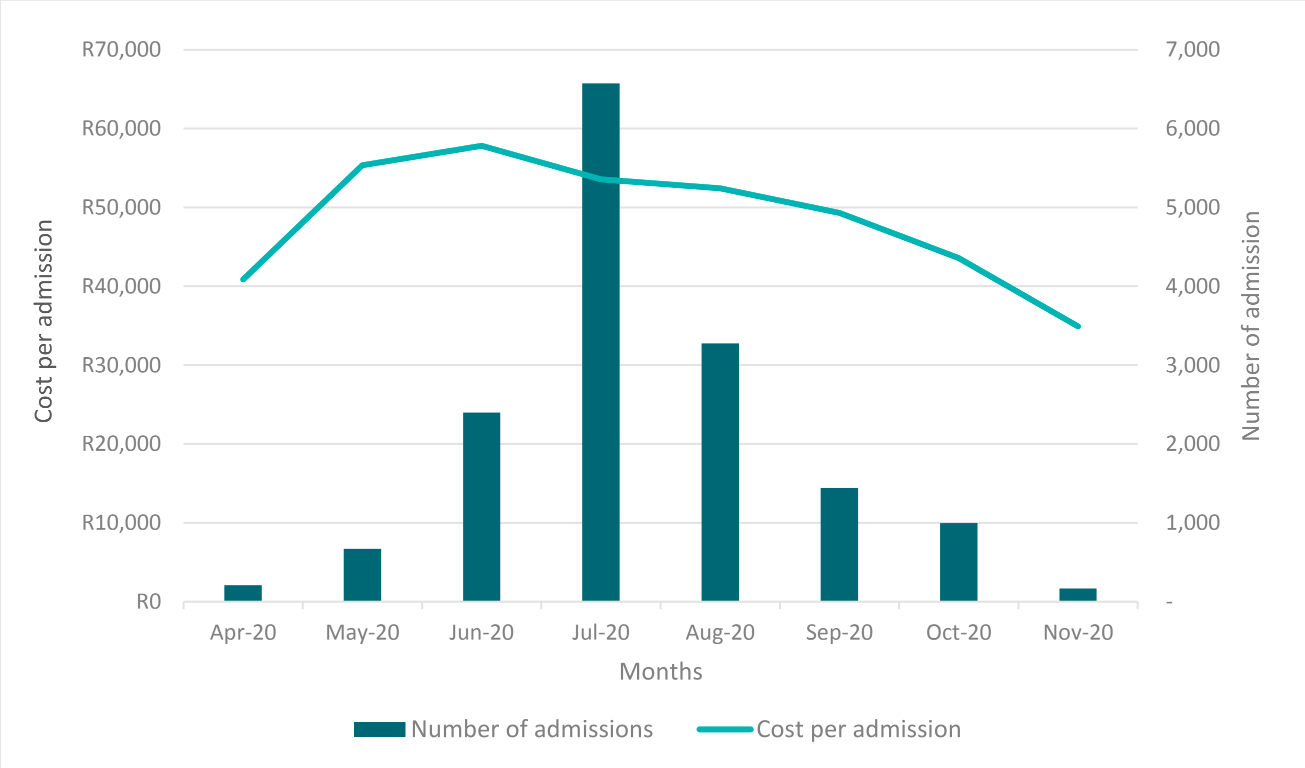
The graph above demonstrates an additional source of cost variability, namely variability by medical scheme. This could be driven by, for example, differing hospital reimbursement and tariff arrangements.

The graph appears to indicate that, as an example, scheme 6 experiences lower costs than scheme 5, even on an age-adjusted basis.

4.4.8. Cost trend by month

The graph below considers the trend in average cost per admission over the course of 2020, by month of admission.

Figure 15: COVID-19 cost per admission and number of admissions over time



The graph above demonstrates a potential downward trend in the cost per admission during the course of 2020. This could be indicative of changing and more efficient treatment approaches as the knowledge of how to treat COVID-19 improved and treatment became more focused.

It could also in part be explained by claims payments incurred, but not yet reported, which would affect the more recent admissions to a greater degree.

Regardless of the exact source of the trend, this points to a further difficulty in using the COVID-19 admission data to develop a reasonable ARM or global fee model that is appropriate for admissions going forward. The majority of the data gathered are for admissions in July and August 2020. Treatment approaches and resulting costs are likely to have developed and improved since then.

That being said, the data available were used to fit optimal predictive models for COVID-19 admission costs and lengths of stay. This is discussed further in the next section.

5. Findings – model fitting

This section of the report provides a summary of the findings in respect of fitting models to the COVID-19 admission data.

A logistic regression model was fitted to the trimmed data. The total benefits paid and the length of stay were log transformed and then fitted to a linear regression model. The log transformation was the most suitable transformation. This approach is similar to other approaches used to model hospital claims as noted in section 2.2 for the reasons given.

The models developed are for the total benefits paid for hospital admissions, the average length of stay and the average daily rate per admission. The method followed to structure and trim the input data was highlighted in section 4 of this report.

The variables considered affecting the three outcomes for the purposes of the models were:

- Age;
- Gender;
- Hospital group;
- Level of care (three levels of care); and
- Existence of a pre-existing condition.

The research considered in the literature review, as well as the data trend analysis performed on trimmed data (Section 4.4) demonstrated the above to be the factors most expected to drive variability in hospital admission severity and cost.

The detailed regression approach and results are discussed further in Annexure 3.

5.1. Risk-adjusted length of stay

The model for length of stay had a good fit, explaining approximately 34% of the variation. The important risk factors were age, level of care, gender and the pre-existing conditions listed below. Hospital groups 3 and 4 showed significant differences from hospital group 1 while hospital group 2 was similar to hospital group 1.

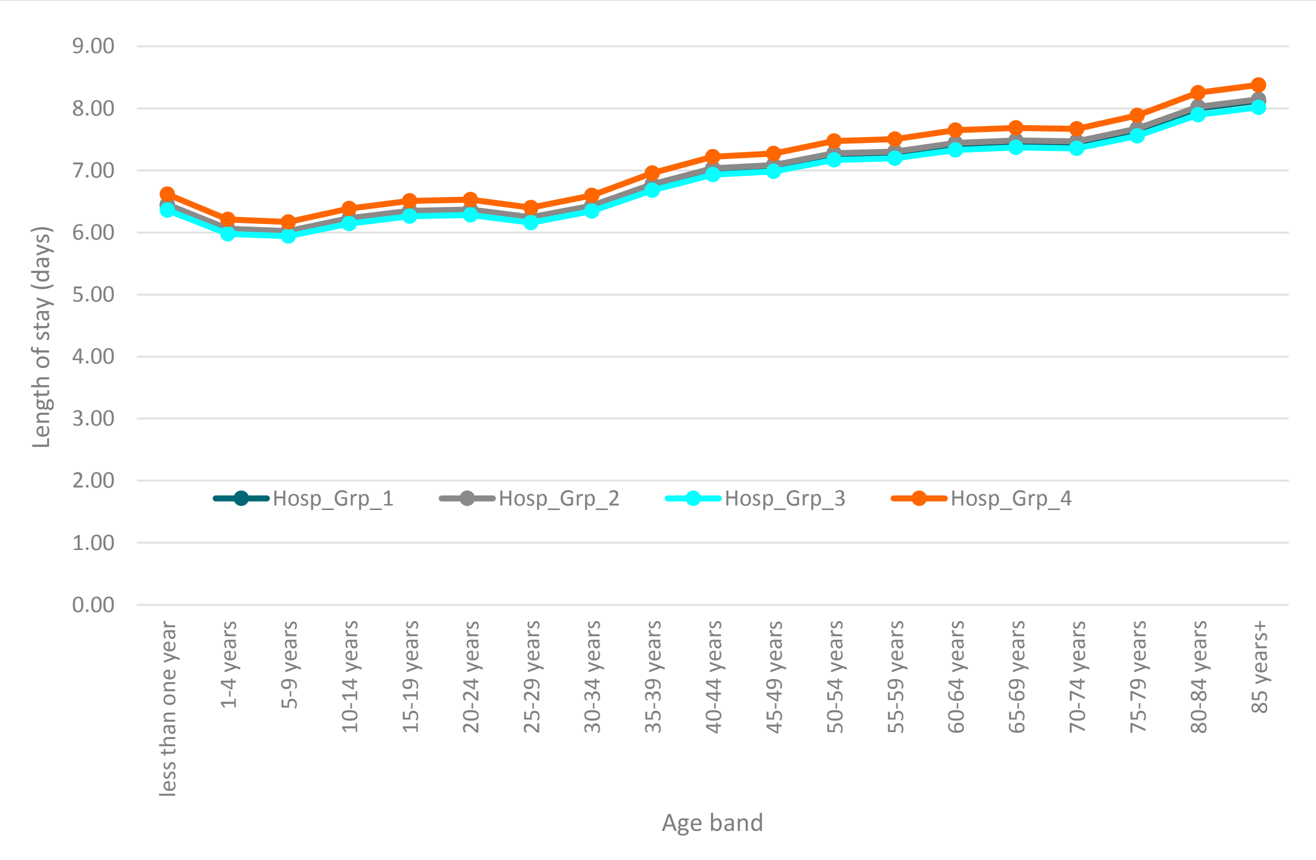
Significant pre-existing conditions: length of stay

- HIV
- Bipolar mood disorder
- Chronic renal disease
- Diabetes mellitus type 1
- Diabetes mellitus type 2
- Epilepsy
- Hyperlipidaemia
- Hypertension
- Hypothyroidism
- Multiple sclerosis
- Tuberculosis

The figure below shows the total benefits paid predicted by the model for COVID-19 admissions adjusted for the risk factors listed above. The results shown are filtered as follows:

- Patients with at least one risk condition
- General ward admissions only
- COVID-19 admissions only
- Females only

Figure 16: Model result: Length of stay by hospital group and age



These results are risk adjusted; there is also minimal variation in the length of stay in the general ward across the four hospital groups. The length of stay increases with age.

There is an online tool showing the model results (with various scenarios) that has been prepared to allow readers to view the model results. This is available on the following link:
https://www.bhfportal.co.za/bhf_global/HOSP/

5.2. Risk-adjusted total benefits paid

The model for total benefits paid (global fee) had a good fit, explaining approximately 54% of the variation. The important risk factors were age, level of care, the hospital group and the pre-existing conditions listed in the table below. Gender does not have a significant impact on the total benefits paid.

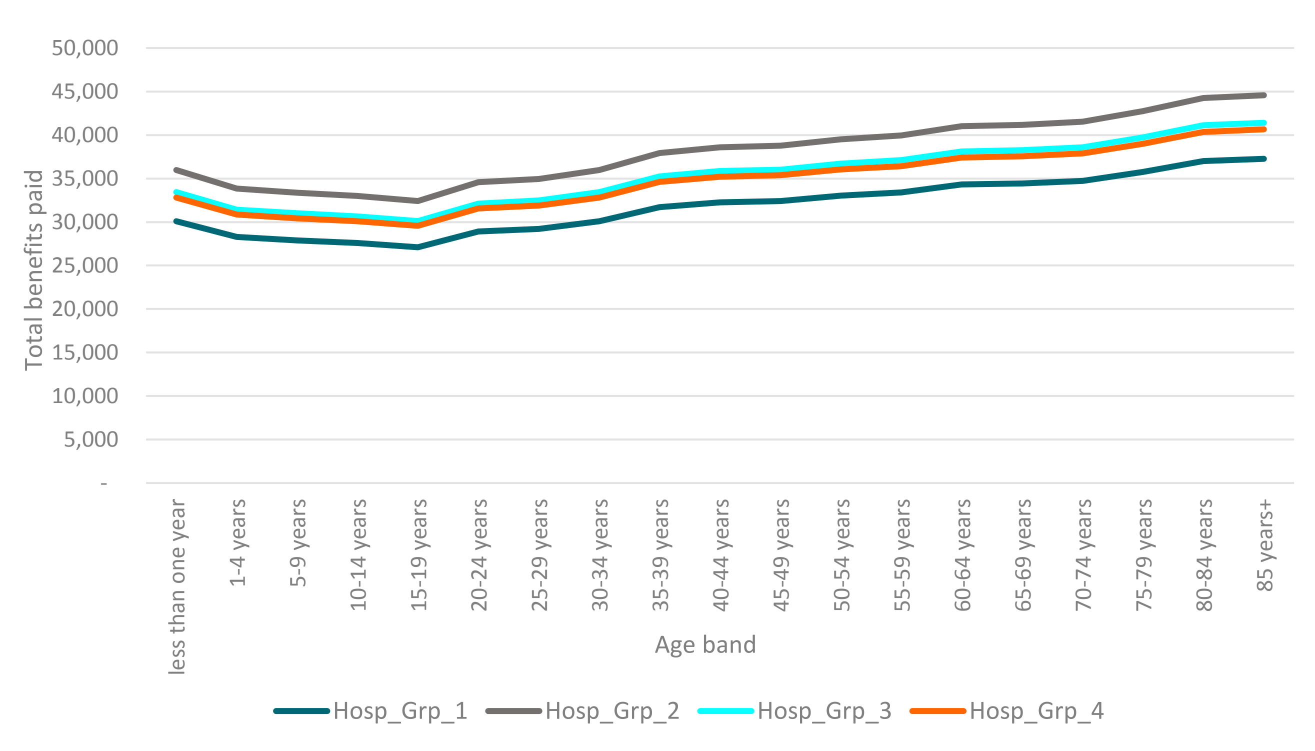
Significant pre-existing conditions: total benefits paid

- HIV
- Bipolar mood disorder
- Chronic renal disease
- Diabetes mellitus type 1
- Diabetes mellitus type 2
- Epilepsy
- Hyperlipidaemia
- Hypertension
- Hypothyroidism
- Multiple sclerosis
- Tuberculosis
- Asthma
- Bronchiectasis
- Cardiac failure
- Chronic obstructive pulmonary disorder
- Parkinson's disease
- Rheumatoid arthritis
- Cancer

The figure below shows the total benefits paid predicted by the model for COVID-19 admissions adjusted for the risk factors listed above. The results shown are filtered as follows:

- Patients with at least one risk condition
- General ward admissions only
- COVID-19 admissions only
- Females only

Figure 17: Model result: Total benefit paid by hospital group and age



These results are risk adjusted; therefore the variation between the total benefit paid by hospital groups is largely due to difference in the tariffs charged and/or negotiated with schemes.

- The total benefits paid increase with age.
- Total benefits paid for hospital groups 3 and 4 are very similar to hospital group 2, being slightly more expensive.
- Hospital group 1 had the lowest total benefits paid while hospital group 2 had the highest total benefits paid across all ages.

5.3. Risk-adjusted average cost per day

The model for average cost per day (*per diem* rate) had a very good fit, explaining approximately 74% of the variation. The important risk factors were age – though not across all 19 age bands, level of care, the hospital group and the pre-existing conditions listed in the table below. Gender does not have a significant impact on the daily benefits paid.

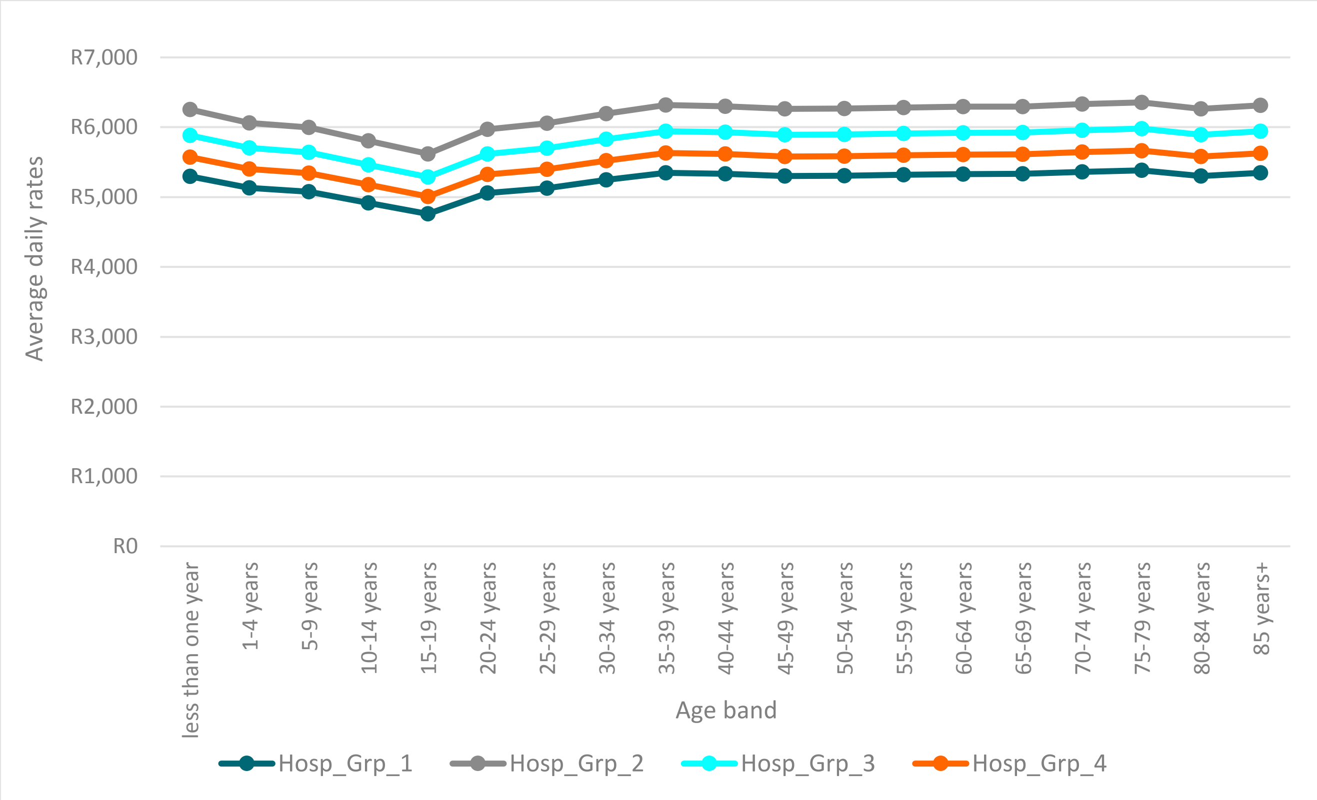
Significant pre-existing conditions: average daily cost

- HIV
- Asthma
- Bronchiectasis
- Diabetes mellitus type 1
- Diabetes mellitus type 2
- Cardiac failure
- Hyperlipidemia
- Hypertension
- Hypothyroidism
- Chronic obstructive pulmonary disorder
- Coronary artery disease
- Rheumatoid arthritis
- Cancer
- Obesity

The figure below shows the total benefits paid predicted by the model for COVID-19 admissions adjusted for the risk factors listed above. The results shown are filtered as follows:

- Patients with at least one risk condition
- General ward admissions only
- COVID-19 admissions only
- Females only

Figure 18: Model result: Average daily rate by hospital group and age



These results are risk adjusted, therefore the variation between the hospital groups is largely due to difference in the tariffs charged and/or negotiated with schemes. The average daily cost in hospital is largely constant across the 19 age bands.

5.4. Goodness of fit of model results

The models for predicting the total benefits paid and the average daily cost per admission were further tested for goodness of fit. While the predictive power of the two models was 54% and 74%, respectively, a further test was carried out.

The model results were compared to the actual observed results to determine the variance of the model from the actual.

The percentage difference was calculated as follows:

$$\%\ difference = \frac{Expected\ Result - Actual\ Result}{Actual\ Result}$$

If the model fit is very good the % difference will be very small, close to zero.

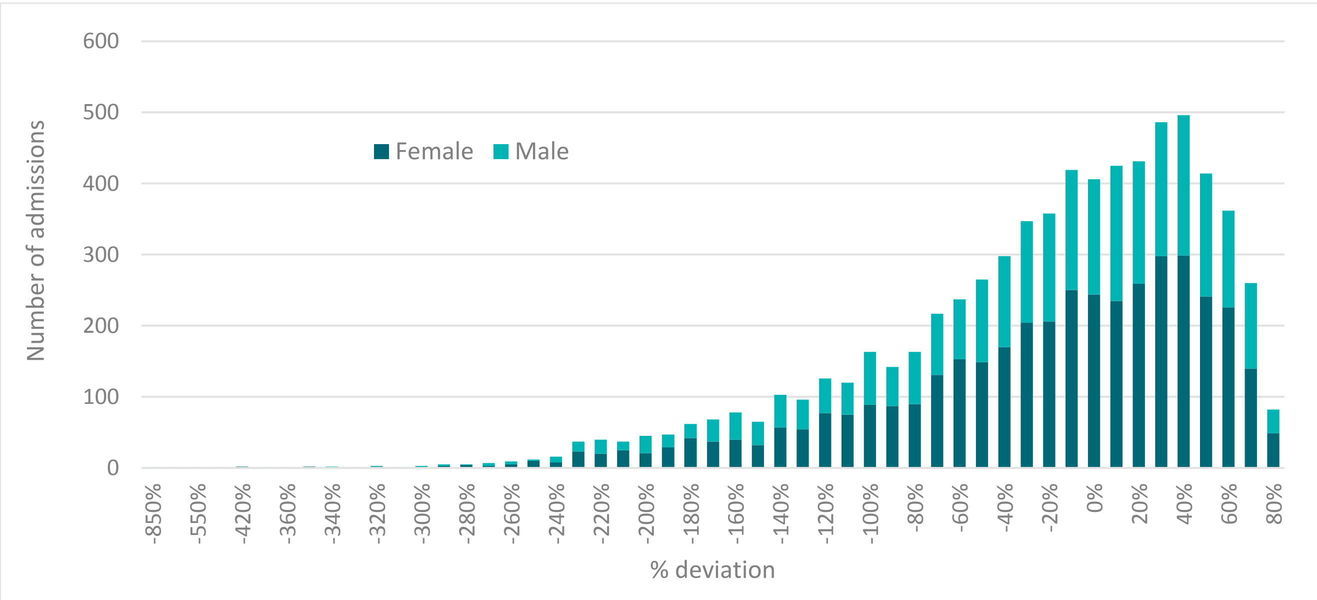
Selected goodness of fit results are shown below. Further results are available in Annexure 4.

5.4.1. Total benefits paid

The predicted total benefits paid were compared to the observed total benefits paid and the following graph shows the percentage differences. The results shown are filtered as follows:

- Patients with at least one risk condition
- General ward admissions only
- COVID-19 admissions only

Figure 19: COVID-19 model goodness of fi: total benefits paid



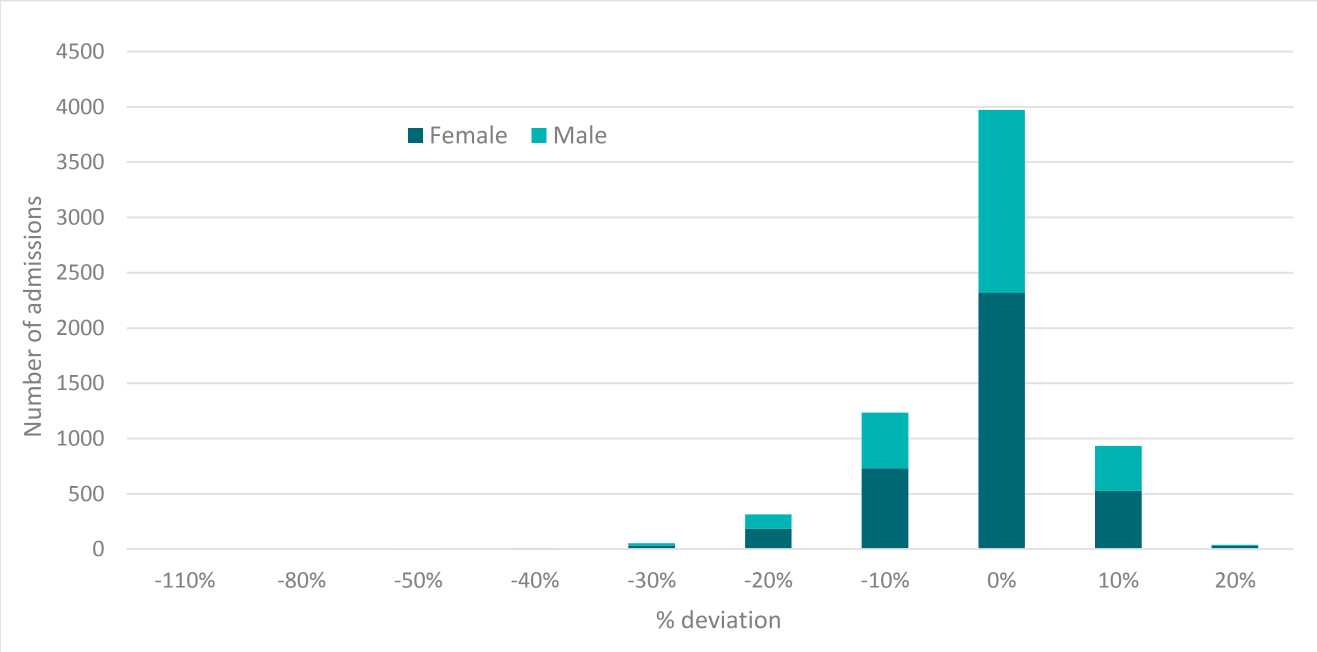
As previously observed, there is a wide disparity in the total benefits paid. This implies that it may be very difficult to implement a global fee arrangement as the appropriate sharing of risk between funders and hospitals will be prone to too much variation.

5.4.2. Average daily costs

The predicted average daily costs were compared to the observed average daily costs and the following shows the percentage differences. The results shown are filtered as follows:

- Patients with at least one risk condition
- General ward admissions only
- COVID-19 admissions only

Figure 20: COVID-19 model goodness of fit; average daily rate



The variation between the model results and the actual results is quite narrow. The bulk of observations show a 10% deviation, with a significant majority falling within a 5% deviation. This implies that a *per diem* model could prove to be practically implementable, noting certain challenges discussed earlier in the document.

6. Conclusions

6.1. Global fee for COVID-19 hospital admissions

The determination of a global fee for COVID-19 is extremely difficult. This is due to a number of factors, including:

- There are limited data, since this condition is new
- There are no standard treatment guidelines on how COVID-19 should be managed in hospital
- Resource constraints due to hospital capacity challenges may have influenced treatment approaches and trends (e.g. with regard to lengths of stay and the level of care)

The available treatment guidelines have also varied over time as new information about the disease has emerged. This adds complexity in terms of determining a global fee for hospital admissions.

Despite these concerns, the model for global fees produced results explaining 54% of observed variation, which is considered relatively positive. However, the goodness of fit testing performed (Section 5.4.1) demonstrated concerns in the spread of observed costs relative to those predicted by the model.

6.2. *Per diem* rate for COVID-19 hospital admissions

From the results shown in Section 5, it appears possible to establish *per diem* rates for COVID-19 hospital admissions. The modelled *per diem* rates quite closely match the observed daily rates for COVID-19 hospital admissions.

Given the modelling outcomes, the daily rates would be varied by level of care, hospital group and existence of a relevant pre-existing condition.

In order to practically implement these daily rates, agreements may be required with hospital groups to establish outlier cases that would be carved out and charged on a fee-for-service basis, where these outliers would be expected to approximately match the trim points employed for limiting the data used for the modelling exercise.

Since admissions progressing to higher levels of care were effectively excluded, these admissions would also presumably carve out and be charged on a fee-for-service basis. The problem with this is that there is a certain amount of discretion with regard to level of care. For example, where the *per diem* rate is low there may be an increased propensity to send the patient to a high care or ICU ward even if not strictly necessary.

Given that hospital group was observed to be a significant factor in determining expected daily costs, it should be noted that this implies the need to negotiate separate daily rates with each hospital group.

6.3. Global fees for other respiratory hospital admissions

From the modelling exercise, it does not appear possible to determine a global fee for other respiratory conditions included in this exercise. While the limitations that exist on COVID-19 highlighted above do not exist for the other known and well-established respiratory conditions, there is still significant variation in terms of goodness of fit. Annexure 9.4 shows the goodness of fit graphs for the other respiratory conditions.

6.4. *Per diem* rate for other respiratory hospital admissions

The same conclusion may be drawn for the *per diem* rate for other respiratory conditions and the *per diem* rate for COVID-19. Perhaps there is more evidence supporting the *per diem* rate than the global fee as the model predicts with greater accuracy the expected daily costs relating to the *per diem* rate. Annexure 9.4 shows the goodness of fit graphs for the other respiratory conditions on the *per diem* model.

6.5. Study limitations

This study had limitations, which are discussed in this section.

6.5.1. Retrospective nature of the study

The analysis supporting this report was done retrospectively. Given the novel nature of COVID-19, this does not mean that it would have been possible to prospectively determine alternative reimbursement rates.

6.5.2. Length of stay in each level of care

Due to the nature of data collected, we were unable to identify the date a patient was moved from one level of care to another. This has an impact in the determination of the global fee and the average daily rate for each admission.

Ultimately for the purpose of the modelling exercise performed, the entire duration of a hospital admission was assumed to occur in the highest level of care observed at any point during the admission.

6.5.3. Claims incurred but not reported

The data were collected for admissions up to 31 December 2020 and extracted shortly after that. There are some claims that may have been incurred but not yet submitted to the schemes, or received but not yet paid. This impacts the model in that certain costs would have been understated given the lack of completeness. This was, however, partly addressed through the use of data trim points, as well as excluding admissions with a total cost under R3 500.

6.5.4. Volume of data

The data incorporated for the purposes of this exercise were in respect of schemes representing 37% of the total scheme industry population. This is considered a relatively large data set. However, the exercise would have benefited from the inclusion of further data. This is particularly relevant given that only six schemes participated, and there is known variability in scheme hospital cost experience, e.g. with relation to differing negotiated reimbursement arrangements.

In addition, there is a possibility that the results were skewed towards a particular population with unique demographic or other characteristics (such as different managed care protocols) that could potentially bias results.

6.5.5. Mortality data

The literature review revealed that mortality is consistently considered to have a significant impact on expected COVID-19 cost experience. The modelling exercise, as well as any ultimately implemented reimbursement structures, would possibly benefit from the consideration of mortality as a predictive variable. However, this was not considered as the mortality data provided were not considered to be complete and/or sufficiently reliable.

7. Recommendations

7.1. Future studies

This is the first of hopefully more reports on ARMs for hospitalisations. The below considers possible extensions to the COVID-19 modelling exercise discussed in this report, as well as additional potential future studies.

7.1.1. COVID-19

The approach used in this report explores expenditure from hospital episodes and identifies appropriate risk factors for modelling hospital-related costs. This approach does not explore the specific interventions in each episode of care.

The team responsible for the bottom-up approach discussed in the introduction to this document is in the process of exploring practical issues with alternative reimbursement of COVID-19 in-hospital costs. The results of that study will be published in due course.

The approach used to model in this report may be improved by including more refined data, including updated data from further cases experienced, as well as getting more schemes to participate in the study.

For the next iterations of data collection, it is also considered key to include the following:

- Outcome of each hospital episode (e.g. mortality)
- Length of stay at each level of care.

7.1.2. Additional conditions

Many of the challenges encountered in modelling predicted hospital costs for COVID-19 (in particular for global fees) are not similarly evident in respect of other conditions and hospital admission categories.

In particular, having considered the data collected in respect of additional respiratory diseases and applying a similar modelling approach, it appears more feasible to determine a global fee for other respiratory conditions. Treatment approaches are more stable, there is a greater depth and history of data and less volatility with regard to cost experience.

There may be benefit in seeking to produce predictive cost models for conditions other than COVID-19 (e.g. respiratory conditions), where the results may prove more stable. This may allow for the initial consideration of an ARM with a greater likelihood of gaining traction as an industry-wide implementable cost model.

Consistent industry-wide ARMs could facilitate competition on efficiency and value as opposed to price.

8. References

1. Vlachos S, Wong A, Metaxa V, Canestrini S, Lopez Soto C, Periselneris J, et al. Hospital Mortality and Resource Implications of Hospitalisation with COVID-19 in London, UK: A Prospective Cohort Study. Crit Care Res Pract [Internet]. 2021;2021(December 2019). Available from: https://pubmed.ncbi.nlm.nih.gov/33564474/

2. Qifang B, Yongsheng W, Shujiang M, Chenfei Y, Xuan Z, Zhen Z, et al. Epidemiology and transmission of COVID-19 in Shenzhen China: Analysis of 391 cases and 1,286 of their close contacts. medRxiv [Internet]. 2020; Available from: https://www.medrxiv.org/content/10.1101/2020.03.03.20028423v3

3. Himmels JPW, Borge TC, Brurberg KG, Gravningen KM, Feruglio SL BJ. COVID-19-EPIDEMIC : COVID-19 and risk factors for hospital admission, severe disease and death – a rapid review, 3rd update [Internet]. COVID-19-EPIDEMIC : COVID-19 and risk factors for hospital admission, severe disease and death – a rapid review, 3rd update. 2020. Available from: https://www.fhi.no/globalassets/dokumenterfiler/rapporter/2020/covid-19-and-risk-factors-for-hospital-admission-severe-disease-and-death-3rd-update-memo-2020-v2.pdf

4. Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J, et al. Clinical Characteristics of 138 Hospitalized Patients with 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China. JAMA - J Am Med Assoc [Internet]. 2020;323(11):1061–9. Available from: https://pubmed.ncbi.nlm.nih.gov/32031570/

5. Holler JG, Eriksson R, Jensen TØ, van Wijhe M, Fischer TK, Søgaard OS, et al. First wave of COVID-19 hospital admissions in Denmark: a Nationwide population-based cohort study. BMC Infect Dis [Internet]. 2021;21(1):1–16. Available from: https://bmcinfectdis.biomedcentral.com/articles/10.1186/s12879-020-05717-w

6. Liu J, Zhang S, Wu Z, Shang Y, Dong X, Li G, et al. Clinical outcomes of COVID-19 in Wuhan, China: a large cohort study. Ann Intensive Care [Internet]. 2020;10(1). Available from: https://doi.org/10.1186/s13613-020-00706-3

7. Cleary SM, Wilkinson T, Tamandjou Tchuem CR, Docrat S, Solanki GC. Cost-effectiveness of intensive care for hospitalized Covid-19 patients: Experience from South Africa. medRxiv [Internet]. 2020;4:1–10. Available from: https://bmchealthservres.biomedcentral.com/track/pdf/10.1186/s12913-021-06081-4.pdf

8. National Department of Health South Africa. Guildelines on Public Private Collaboration in response to COVID-19. 2020.

9. Chen J, Gu Y, Jones AM, Peng B. Modelling Healthcare Costs: A Semiparametric Extension of Generalised Linear Models. SSRN Electron J [Internet]. 2020;(740654). Available from: https://ideas.repec.org/p/yor/hectdg/20-03.html

10. Borislava Mihaylova, Andrew Briggsb, Anthony O'Hagan Sgt. Review of Statistical Methods For Analysing Healthcare Resources and Costs. 2010;1131(20):1127–31. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/hec.1653

11. Gregori D, Petrinco M, Bo S, Desideri A, Merletti F, Pagano E. Regression models for analyzing costs and their determinants in health care: An introductory review. Int J Qual Heal Care [Internet]. 2011;23(3):331–41. Available from: https://pubmed.ncbi.nlm.nih.gov/21504959/

12. Malehi AS, Pourmotahari F, Angali KA. Statistical models for the analysis of skewed healthcare cost data: a simulation study. Health Econ Rev [Internet]. 2015;5(1). Available from: https://healtheconomicsreview.biomedcentral.com/track/pdf/10.1186/s13561-015-0045-7.pdf

13. The Independent Hospital Pricing Authority. National Pricing Model Technical Specifications 2013-2014 Version 1.0. 2013;1–59. Available from: http://www.ihpa.gov.au/internet/ihpa/publishing.nsf/Content/CA25794400122452CA257B15007D2F51/$File/National-Pricing-Model-Technical-specifications-2013-14.pdf

14. Jackson T. Cost estimates for hospital inpatient care in Australia: Evaluation of alternative sources. Aust N Z J Public Health [Internet]. 2000;24(3):234–41. Available from: https://pubmed.ncbi.nlm.nih.gov/10937398/

15. Medarevic AP. Describing serbian hospital activity using Australian refined diagnosis related groups: A case study in Vojvodina Province. Zdr Varst. 2020;59(1):18–26.

16. Niangoran AK, Diako DJ, Mensah EP, Achiepo OYM. Method for automatically processing outliers of a quantitative variable. Int J Adv Comput Sci Appl [Internet]. 2020;11(7):407–11. Available from: https://thesai.org/Downloads/Volume11No7/Paper_53-Method_for_Automatically_Processing_Outliers.pdf

17. Church RM. How To Look At Data: a Review of John W. Tukey'S Exploratory Data Analysis 1. J Exp Anal Behav [Internet]. 1979;31(3):433–40. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1332871/

18. Dias SS, Martins MFO. HIV AIDS Length of Stay Outliers. Procedia Comput Sci [Internet]. 2015;64:984–92. Available from: http://dx.doi.org/10.1016/j.procs.2015.08.617

19. Jain V. A Data-driven and cost-effective approach towards Risk Based Monitoring. :1–7. Available from: https://www.lexjansen.com/phuse/2019/ar/AR02.pdf

9. Annexures

9.1. ANNEXURE 1: Identification of Respiratory Admissions

Diagnosis	Diagnosis description
U07.1	COVID-19, virus identified
U07.2	COVID-19, virus not identified
J12.8	Other viral pneumonia
J12.9	Viral pneumonia, unspecified
J18.0	Bronchopneumonia, unspecified
J18.1	Lobar pneumonia, unspecified
J18.2	Hypostatic pneumonia, unspecified
J18.8	Other pneumonia, organism unspecified
J18.9	Pneumonia, unspecified
J22	Unspecified acute lower respiratory infection
J80	ARDS
J96.0	Acute respiratory failure
J96.9	Respiratory failure, unspecified
J45	Asthma
J45.0	Predominantly allergic asthma
J45.1	Nonallergic asthma
J45.8	Mixed asthma
J45.9	Asthma, unspecified
J46	Status asthmaticus
J43	Emphysema
J43.0	Macleod's syndrome
J43.1	Panlobular emphysema
J43.2	Centrilobular emphysema
J43.8	Other emphysema
J43.9	Emphysema, unspecified
J44	Other chronic obstructive pulmonary disease
J44.0	Chronic obstructive pulmonary disease with acute lower respiratory infection
J44.1	Chronic obstructive pulmonary disease with acute exacerbation, unspecified
J44.8	Other specified chronic obstructive pulmonary disease
J44.9	Chronic obstructive pulmonary disease, unspecified
J20	Acute bronchitis
J20.0	Acute bronchitis due to *Mycoplasma pneumoniae*
J20.1	Acute bronchitis due to *Haemophilus influenzae*
J20.2	Acute bronchitis due to streptococcus
J20.3	Acute bronchitis due to coxsackievirus
J20.4	Acute bronchitis due to parainfluenza virus
J20.5	Acute bronchitis due to respiratory syncytial virus
J20.6	Acute bronchitis due to rhinovirus
J20.7	Acute bronchitis due to echovirus
J20.8	Acute bronchitis due to other specified organisms
J20.9	Acute bronchitis, unspecified
J40	Bronchitis, not specified as acute or chronic

9.2. ANNEXURE 2: Data Specification

Hospital admission data were requested from medical schemes that agreed to participate and submit data.

Two data sets were requested, namely a COVID-19 data set and a data set for other conditions. The COVID-19 data set included all hospital admissions with a primary or secondary discharge ICD-10 diagnosis code of U07.1 or U07.2.

The other conditions included key respiratory disorders. Data were requested for anyone with a primary or secondary discharge diagnosis as per the second table in Annexure 1 above.

Ultimately data were requested for all COVID-19 claims between 1 January 2020 and 31 December 2020, and for all other respiratory conditions between 1 January 2019 and 31 December 2020.

Within each data set, three data tables were requested, as described further below.

DATA TABLE	DESCRIPTION OF DATA REQUESTED
1. Hospital event level detail	Data providing further information per hospital admission, including, for example, admission dates, discharge dates, diagnosis codes and procedure codes
2. Claim line level detail	Detailed claim line level data per hospital admissions. This included relevant dates, amounts and codes providing additional descriptions of the service provided and who provided it
3. Demographic data	Demographic information per patient including date of birth and gender, as well as information on risk factors such as chronic diseases

Unique identifying codes were created per patient and per hospital admission for the linking of data tables. These were also created in a manner that ensured all data remained anonymised and personal patient information was kept confidential.

9.3. ANNEXURE 3: Model Construction

Logistic regression analysis was used to identify risk factors for hospital-related expenditure and the length of stay in hospital. Three models were constructed for total hospital benefits, length of stay and average daily rate.

For each of these models, logistic regression analysis was performed as well as for each explanatory variable; a variable was considered significant at probability value (p-value) < 0.05. The variables considered affecting the three outcomes were as follows:

- Age (19 age bands used);
- Gender (male or female);
- Admission reason (the 6 respiratory conditions);
- Hospital group (classified into four groupings);
- Admission year (2019 or 2020);
- Level of care (3 levels of care); and
- Pre-existing conditions (29 conditions listed in the data specification)

For the model to be more useful, it is necessary to reduce the number of variables. In the case of pre-existing conditions, these were grouped into one class reflecting that an individual has at least one condition to be classified as having a risky pre-existing medical condition. This is to allow for simpler application of the model as the number of indicators for pre-existing conditions reduces from 29 to only one. The final model therefore had only one binary indicator for the pre-existing conditions.

ANNEXURE 3.1. Hospital length of stay

The average length of stay was modelled using the above approach. The results from the initial model are highlighted in the table below.

The model can explain approximately 33% of the observed outcome and may be relied on (statistically significant with p-value <0.001).

The statistically significant risk factors are as follows:

- Age – though there is no significant difference between age ranges 50-54 years and 55-59
- Gender (p-value = 0.320)
- Admission reason
- Level of care
- Hospital network - though there is no significant difference between hospital group 1 and hospital group 2 (p-value = 0.259)

Pre-existing conditions that were significant were as follows: HIV; bipolar mood disorder; chronic renal disease; diabetes mellitus type 1; diabetes mellitus type 2; epilepsy; hyperlipidaemia; hypertension; hypothyroidism; multiple sclerosis and tuberculosis.

For the purposes of constructing the final model, any patient having any one of these was classified as having a risk condition.

Source	SS	df	MS			
				Number of obs	=	84,650
				F(58, 84591)	=	749.94
Model	317.085454	58	5.46699059	Prob > F	=	0.0000
Residual	616.660761	84,591	.00728991	R-squared	=	0.3396
				Adj R-squared	=	0.3391
Total	933.746215	84,649	.0110308	Root MSE	=	.08538

T_LOS	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
AgeBand						
less than one year	.0230073	.0017141	13.42	0.000	.0196477	.0263669
1-4 years	.0357348	.0013718	26.05	0.000	.0330461	.0384236
5-9 years	.0375615	.0016089	23.35	0.000	.0344079	.040715
10-14 years	.0309573	.0018851	16.42	0.000	.0272625	.0346521
15-19 years	.0275165	.0019683	13.98	0.000	.0236586	.0313744
20-24 years	.0265497	.0022459	11.82	0.000	.0221477	.0309516
25-29 years	.0299987	.0022218	13.50	0.000	.025644	.0343534
30-34 years	.0249012	.0018026	13.81	0.000	.0213682	.0284342
35-39 years	.0149598	.0016315	9.17	0.000	.0117622	.0181574
40-44 years	.0077951	.0015691	4.97	0.000	.0047197	.0108704
45-49 years	.0055786	.0014679	3.80	0.000	.0027016	.0084556
55-59 years	-.0014428	.0014508	-0.99	0.320	-.0042863	.0014007
60-64 years	-.0059984	.0016008	-3.75	0.000	-.0091358	-.0028609
65-69 years	-.0076124	.0017811	-4.27	0.000	-.0111034	-.0041214
70-74 years	-.0087298	.001885	-4.63	0.000	-.0124244	-.0050351
75-79 years	-.0150878	.0020213	-7.46	0.000	-.0190495	-.011126
80-84 years	-.0226383	.0022425	-10.09	0.000	-.0270336	-.0182429
85 years+	-.0259339	.0023034	-11.26	0.000	-.0304486	-.0214191
Sex						
M	.0045924	.0006186	7.42	0.000	.0033799	.0058048
Admin_Reason						
Bronchitis	-.0202863	.001264	-16.05	0.000	-.0227638	-.0178088
COPD	-.0644422	.0015866	-40.62	0.000	-.0675519	-.0613325
Covid	-.1232656	.0010557	-116.76	0.000	-.1253348	-.1211965
Pneumonia	-.0611114	.0008605	-71.01	0.000	-.062798	-.0594247
Respiratory_Failure	-.184204	.0019783	-93.11	0.000	-.1880814	-.1803266
1.hiv	-.0176777	.0014474	-12.21	0.000	-.0205146	-.0148408
1.ads	.0113819	.0128228	0.89	0.375	-.0137507	.0365144
1.ast	.0004334	.0008548	0.51	0.612	-.0012419	.0021088
1.bmd	-.0082744	.0021118	-3.92	0.000	-.0124134	-.0041354
1.bce	-.002228	.0043237	-0.52	0.606	-.0107025	.0062464
1.chf	-.0028084	.0081343	-0.35	0.730	-.0187517	.0131348
1.cmy	-.0026075	.0016021	-1.63	0.104	-.0057476	.0005326
1.cop	-.0028837	.0020168	-1.43	0.153	-.0068365	.0010692
1.crf	-.0071626	.0027812	-2.58	0.010	-.0126137	-.0017115
1.cad	.0017222	.0025435	0.68	0.498	-.0032631	.0067075
1.chd	.0499105	.0493034	1.01	0.311	-.0467237	.1465447
1.db1	-.0470996	.0382336	-1.23	0.218	-.1220371	.0278379
1.dm1	-.0081861	.0028805	-2.84	0.004	-.0138319	-.0025402
1.dm2	-.0105399	.0011072	-9.52	0.000	-.0127099	-.0083699
1.dys	-.0017459	.0022985	-0.76	0.448	-.006251	.0027592
1.epl	-.0103126	.0018075	-5.71	0.000	-.0138553	-.0067698
1.glc	-.0005415	.002587	-0.21	0.834	-.005612	.0045291
1.hae	-.0087485	.0349086	-0.25	0.802	-.077169	.059672
1.hyl	.0084685	.0011365	7.45	0.000	.0062409	.0106961
1.hyp	-.0058065	.0009342	-6.22	0.000	-.0076375	-.0039756
1.hyt	.0429327	.0137289	3.13	0.002	.0160242	.0698412
1.mss	-.0284364	.0137072	-2.07	0.038	-.0553025	-.0015704
1.par	-.0046951	.0044277	-1.06	0.289	-.0133733	.0039831
1.rha	-.0024995	.0021742	-1.15	0.250	-.0067609	.0017619
1.scz	-.0015394	.0060266	-0.26	0.798	-.0133515	.0102726
1.sle	-.0062723	.004903	-1.28	0.201	-.0158822	.0033375
0.uci	0	(omitted)				
1.cancerid	.0030094	.0021509	1.40	0.162	-.0012064	.0072252
1.obesityid	-.0018042	.0114281	-0.16	0.875	-.0242032	.0205948
1.tbid	-.0313875	.0018318	-17.13	0.000	-.0349779	-.0277971
Care_Level						
High Care	-.0252849	.0012036	-21.01	0.000	-.027644	-.0229257
ICU	-.0245882	.0018535	-13.27	0.000	-.028221	-.0209553
Net_work						
Hosp_Grp_2	-.0009189	.0008145	-1.13	0.259	-.0025153	.0006775
Hosp_Grp_3	.0020038	.0007976	2.51	0.012	.0004404	.0035671
Hosp_Grp_4	-.0055829	.0008242	-6.77	0.000	-.0071983	-.0039675
_cons	.5044833	.0013825	364.91	0.000	.5017737	.507193

The final model for length of stay was as follows:

Source	SS	df	MS
Model	309.060405	32	9.65813766
Residual	607.784603	84,031	.007232862
Total	916.845008	84,063	.010906642

Number of obs	=	84,064
F(32, 84031)	=	1335.31
Prob > F	=	0.0000
R-squared	=	0.3371
Adj R-squared	=	0.3368
Root MSE	=	.08505

T_LOS	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
AgeBand						
less than one year	.0229293	.0017275	13.27	0.000	.0195434	.0263153
1-4 years	.0354471	.0013832	25.63	0.000	.0327361	.0381581
5-9 years	.0367321	.0016103	22.81	0.000	.0335759	.0398882
10-14 years	.029881	.0018819	15.88	0.000	.0261924	.0335695
15-19 years	.0261288	.0019608	13.33	0.000	.0222857	.0299719
20-24 years	.0254206	.0022392	11.35	0.000	.0210319	.0298093
25-29 years	.0293398	.0022175	13.23	0.000	.0249936	.033686
30-34 years	.0235352	.0017919	13.13	0.000	.0200231	.0270472
35-39 years	.0133269	.0016178	8.24	0.000	.010156	.0164978
40-44 years	.0062806	.0015605	4.02	0.000	.0032221	.0093391
45-49 years	.004915	.0014658	3.35	0.001	.0020421	.007788
55-59 years	-.0007695	.0014498	-0.53	0.596	-.003611	.0020721
60-64 years	-.0042256	.0015918	-2.65	0.008	-.0073455	-.0011056
65-69 years	-.0051386	.0017626	-2.92	0.004	-.0085932	-.001684
70-74 years	-.0047795	.0018536	-2.58	0.010	-.0084125	-.0011464
75-79 years	-.0097952	.0019773	-4.95	0.000	-.0136707	-.0059198
80-84 years	-.0177121	.002194	-8.07	0.000	-.0220122	-.013412
85 years+	-.0203653	.0022516	-9.04	0.000	-.0247784	-.0159521
Sex						
M	.0044548	.0006128	7.27	0.000	.0032538	.0056558
Rk_CDL						
No	.0169506	.0011316	14.98	0.000	.0147326	.0191686
Care_Level						
High Care	-.0249251	.0012107	-20.59	0.000	-.0272981	-.0225522
ICU	-.0208874	.0018815	-11.10	0.000	-.0245752	-.0171996
Admin_Reason						
Bronchitis	-.0197083	.0012471	-15.80	0.000	-.0221527	-.017264
COPD	-.0645941	.0014919	-43.30	0.000	-.0675183	-.06167
Covid	-.1198499	.0011083	-108.14	0.000	-.1220222	-.1176777
Pneumonia	-.06129	.0008312	-73.73	0.000	-.0629192	-.0596608
Respiratory_Failure	-.2004569	.0020512	-97.73	0.000	-.2044771	-.1964366
Net_work						
Hosp_Grp_2	-.0005261	.0008139	-0.65	0.518	-.0021214	.0010693
Hosp_Grp_3	.0022755	.0007957	2.86	0.004	.000716	.003835
Hosp_Grp_4	-.0054176	.0008235	-6.58	0.000	-.0070317	-.0038036
Year						
2020	-.0004212	.0006957	-0.61	0.545	-.0017849	.0009424
CDL_Count	.0042058	.0010483	4.01	0.000	.002151	.0062606
_cons	.4873332	.0018115	269.02	0.000	.4837826	.4908837

Similar conclusions may be drawn on the strength of the model (R^2 = 33,7% and p-value<0.000). The expected lengths of stay presented in the report were constructed using this model.

ANNEXURE 3.2. Total hospital benefits paid

The total in hospital benefits paid was modelled using the approach described in Appendix 2. The results from the initial model are highlighted in the table below.

Source	SS	df	MS			
				Number of obs	=	84,650
				F(58, 84591)	=	1720.09
Model	26886.253	58	463.556086	Prob > F	=	0.0000
Residual	22796.9089	84,591	.269495678	R-squared	=	0.5412
				Adj R-squared	=	0.5408
Total	49683.1619	84,649	.586931468	Root MSE	=	.51913

T_TPD	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
AgeBand						
less than one year	-.102699	.0104219	-9.85	0.000	-.1231259	-.0822721
1-4 years	-.1654703	.0083409	-19.84	0.000	-.1818185	-.1491221
5-9 years	-.1780611	.0097826	-18.20	0.000	-.197235	-.1588872
10-14 years	-.188993	.0114618	-16.49	0.000	-.211458	-.166528
15-19 years	-.2072688	.0119677	-17.32	0.000	-.2307253	-.1838123
20-24 years	-.1435684	.0136554	-10.51	0.000	-.1703329	-.1168038
25-29 years	-.1312309	.0135088	-9.71	0.000	-.1577081	-.1047538
30-34 years	-.1061294	.0109599	-9.68	0.000	-.1276109	-.084648
35-39 years	-.0553543	.0099195	-5.58	0.000	-.0747965	-.0359122
40-44 years	-.0359152	.0095402	-3.76	0.000	-.0546139	-.0172165
45-49 years	-.0234309	.0089248	-2.63	0.009	-.0409235	-.0059384
55-59 years	.016606	.008821	1.88	0.060	-.000683	.033895
60-64 years	.0486987	.0097328	5.00	0.000	.0296225	.067775
65-69 years	.0530385	.0108295	4.90	0.000	.0318127	.0742643
70-74 years	.0625523	.0114614	5.46	0.000	.0400881	.0850165
75-79 years	.0957837	.0122899	7.79	0.000	.0716956	.1198719
80-84 years	.1307371	.013635	9.59	0.000	.1040126	.1574617
85 years+	.1385622	.0140053	9.89	0.000	.1111119	.1660126
Sex						
M	-.0065128	.0037613	-1.73	0.083	-.0138849	.0008593
Admin_Reason						
Bronchitis	-.0596458	.0076855	-7.76	0.000	-.0747093	-.0445824
COPD	.2363571	.0096466	24.50	0.000	.2174499	.2552643
Covid	.8953387	.0064187	139.49	0.000	.8827581	.9079194
Pneumonia	.1706544	.0052323	32.62	0.000	.1603992	.1809096
Respiratory_Failure	1.341225	.0120282	111.51	0.000	1.31765	1.3648
1.hiv	.1029258	.0088005	11.70	0.000	.0856768	.1201747
1.ads	-.0788679	.0779645	-1.01	0.312	-.2316777	.0739419
1.ast	.0133448	.0051971	2.57	0.010	.0031585	.0235312
1.bmd	.0416267	.0128398	3.24	0.001	.0164607	.0667927
1.bce	.0592251	.0262889	2.25	0.024	.007699	.1107511
1.chf	.2015377	.0494581	4.07	0.000	.1046002	.2984751
1.cmy	-.0061363	.0097409	-0.63	0.529	-.0252285	.0129558
1.cop	.0378856	.0122623	3.09	0.002	.0138516	.0619196
1.crf	.0727787	.0169099	4.30	0.000	.0396354	.105922
1.cad	-.026417	.0154651	-1.71	0.088	-.0567285	.0038945
1.chd	-.1368482	.2997721	-0.46	0.648	-.7243993	.4507028
1.db1	.2840714	.2324662	1.22	0.222	-.1715604	.7397032
1.dm1	.0646219	.0175141	3.69	0.000	.0302945	.0989494
1.dm2	.0231169	.0067317	3.43	0.001	.0099229	.0363108
1.dys	.0260877	.0139753	1.87	0.062	-.0013039	.0534792
1.epl	.0340885	.0109901	3.10	0.002	.012548	.055629
1.glc	.0109687	.0157296	0.70	0.486	-.0198611	.0417986
1.hae	.0389035	.2122496	0.18	0.855	-.377104	.4549109
1.hyl	-.0411122	.0069104	-5.95	0.000	-.0546564	-.0275679
1.hyp	-.0173105	.00568	-3.05	0.002	-.0284432	-.0061778
1.hyt	-.0854298	.0834738	-1.02	0.306	-.2490377	.0781781
1.mss	.1622547	.0833421	1.95	0.052	-.0010951	.3256044
1.par	.0402828	.0269208	1.50	0.135	-.0124819	.0930474
1.rha	.0314217	.0132194	2.38	0.017	.0055118	.0573316
1.scz	-.0526544	.0366426	-1.44	0.151	-.1244737	.0191648
1.sle	.0207325	.029811	0.70	0.487	-.0376968	.0791618
0.uci	0	(omitted)				
1.cancerid	.057001	.0130779	4.36	0.000	.0313684	.0826337
1.obesityid	.1438041	.0694847	2.07	0.038	.0076146	.2799935
1.tbid	.0983886	.0111378	8.83	0.000	.0765585	.1202187
Care_Level						
High Care	.6070495	.0073183	82.95	0.000	.5927057	.6213934
ICU	1.05972	.0112697	94.03	0.000	1.037632	1.081809
Net_work						
Hosp_Grp_2	.1791037	.0049522	36.17	0.000	.1693975	.1888099
Hosp_Grp_3	.1089801	.0048497	22.47	0.000	.0994747	.1184855
Hosp_Grp_4	.0843023	.0050111	16.82	0.000	.0744805	.094124
_cons	9.493945	.0084058	1129.46	0.000	9.47747	9.51042

The model for total hospital benefits paid explains approximately 54% of the observed outcome and maybe relied on (statistically significant with p-value <0.000).

The statistically significant risk factors are as follows:

- Age – though there is no significant difference between age ranges 50-54 years and 55-59
- Admission reason
- Level of care
- Hospital network

Gender was not statistically different for the total hospital benefits paid (p-value = 0.083); therefore the total hospital benefits paid for males and females may be treated as similar.

Pre-existing conditions that were significant are: HIV; asthma; bipolar mood disorder; bronchiectasis; cardiac failure; chronic obstructive pulmonary disorder; chronic renal disease; diabetes mellitus type 1; diabetes mellitus type 2; epilepsy; hyperlipidaemia; hypertension; multiple sclerosis; Parkinson's disease; rheumatoid arthritis; cancer and tuberculosis.

Similarly, for purposes of constructing the final model, any patient having any one of these was classified as having a risk condition.

Therefore, the final model for total hospital benefits paid was as follows:

Source	SS	df	MS			
Model	26866.7659	32	839.586435	Number of obs	=	84,650
Residual	22816.3959	84,617	.269643168	F(32, 84617)	=	3113.69
				Prob > F	=	0.0000
				R-squared	=	0.5408
				Adj R-squared	=	0.5406
Total	49683.1619	84,649	.586931468	Root MSE	=	.51927

T_TPD	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
AgeBand						
less than one year	-.0936104	.0104592	-8.95	0.000	-.1141103	-.0731105
1-4 years	-.1554992	.0083149	-18.70	0.000	-.1717963	-.1392021
5-9 years	-.1691494	.0096559	-17.52	0.000	-.1880748	-.150224
10-14 years	-.1806049	.0113415	-15.92	0.000	-.2028341	-.1583758
15-19 years	-.1982225	.0118644	-16.71	0.000	-.2214766	-.1749684
20-24 years	-.1334669	.0135689	-9.84	0.000	-.1600617	-.106872
25-29 years	-.1227264	.0134432	-9.13	0.000	-.149075	-.0963778
30-34 years	-.0939213	.0108696	-8.64	0.000	-.1152257	-.0726169
35-39 years	-.0409591	.0098176	-4.17	0.000	-.0602015	-.0217167
40-44 years	-.0237655	.0094759	-2.51	0.012	-.0423382	-.0051927
45-49 years	-.0189309	.0089094	-2.12	0.034	-.0363933	-.0014685
55-59 years	.0106549	.008807	1.21	0.226	-.0066067	.0279166
60-64 years	.0374888	.0096621	3.88	0.000	.0185511	.0564265
65-69 years	.0410824	.0106958	3.84	0.000	.0201188	.0620461
70-74 years	.0494301	.0112437	4.40	0.000	.0273924	.0714677
75-79 years	.0786183	.011988	6.56	0.000	.0551219	.1021148
80-84 years	.1132573	.0133007	8.52	0.000	.0871879	.1393267
85 years+	.1202417	.0136699	8.80	0.000	.0934488	.1470345
Sex						
M	-.0045327	.0037273	-1.22	0.224	-.0118383	.0027728
Rk_CDL						
No	-.0247845	.0255537	-0.97	0.332	-.0748696	.0253007
Care_Level						
High Care	.6028586	.007312	82.45	0.000	.5885271	.6171901
ICU	1.053694	.0112481	93.68	0.000	1.031648	1.07574
Admin_Reason						
Bronchitis	-.0543142	.007536	-7.21	0.000	-.0690847	-.0395436
COPD	.2499801	.0090677	27.57	0.000	.2322074	.2677527
Covid	.8466395	.0066353	127.60	0.000	.8336344	.8596446
Pneumonia	.1731978	.0049401	35.06	0.000	.1635152	.1828804
Respiratory_Failure	1.347637	.0118483	113.74	0.000	1.324414	1.37086
Net_work						
Hosp_Grp_2	.1789867	.00495	36.16	0.000	.1692847	.1886887
Hosp_Grp_3	.1053309	.0048452	21.74	0.000	.0958344	.1148273
Hosp_Grp_4	.0870936	.0050106	17.38	0.000	.0772728	.0969143
Year						
2020	.0737061	.0042371	17.40	0.000	.0654014	.0820108
CDL_Count	-.0022529	.0256582	-0.09	0.930	-.0525429	.048037
_cons	9.487491	.0269492	352.05	0.000	9.434671	9.540312

The same conclusions drawn on the total hospital benefits paid may be drawn on the strength of the model (R^2 = 54% and p-value<0.000), the only difference being the risk pre-existing condition indicator, which is not statistically significant (p-value = 0.332).

ANNEXURE 3.3. Average daily hospital benefit

The average daily hospital benefits paid were modelled using the approach described in Appendix 2. The only difference was total benefits hospital benefits paid was the response variable while length of stay was included as an explanatory variable in the model construction. The average daily hospital benefits paid was calculated from the final model.

The results from the initial model are highlighted in the table below.

Source	SS	df	MS			
				Number of obs	=	84,650
				F(59, 84590)	=	4168.52
Model	36968.2308	59	626.580183	Prob > F	=	0.0000
Residual	12714.9311	84,590	.150312461	R-squared	=	0.7441
				Adj R-squared	=	0.7439
Total	49683.1619	84,649	.586931468	Root MSE	=	.3877

T_TPD	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
LOS	.1188125	.0004588	258.99	0.000	.1179133	.1197117
AgeBand						
less than one year	-.0240711	.0077893	-3.09	0.002	-.0393381	-.008804
1-4 years	-.0586296	.0062429	-9.39	0.000	-.0708657	-.0463936
5-9 years	-.0702711	.0073178	-9.60	0.000	-.084614	-.0559282
10-14 years	-.0992458	.008567	-11.58	0.000	-.116037	-.0824545
15-19 years	-.1275482	.0089431	-14.26	0.000	-.1450766	-.1100198
20-24 years	-.0682334	.0102024	-6.69	0.000	-.0882301	-.0482367
25-29 years	-.0529934	.0100933	-5.25	0.000	-.0727762	-.0332106
30-34 years	-.0280563	.0081908	-3.43	0.001	-.0441102	-.0120025
35-39 years	-.0069146	.0074105	-0.93	0.351	-.0214393	.00761
40-44 years	-.0041349	.0071259	-0.58	0.562	-.0181017	.0098319
45-49 years	-.0046092	.0066657	-0.69	0.489	-.0176739	.0084556
55-59 years	.0077465	.0065879	1.18	0.240	-.0051656	.0206587
60-64 years	.0115481	.0072702	1.59	0.112	-.0027013	.0257976
65-69 years	.0088129	.0080896	1.09	0.276	-.0070427	.0246685
70-74 years	.0113913	.008562	1.33	0.183	-.0053901	.0281727
75-79 years	.0146941	.0091838	1.60	0.110	-.0033061	.0326943
80-84 years	.0040957	.0101948	0.40	0.688	-.015886	.0240773
85 years+	.0086395	.0104716	0.83	0.409	-.0118848	.0291638
Sex						
M	.0027957	.0028093	1.00	0.320	-.0027104	.0083018
Admin_Reason						
Bronchitis	-.0764595	.0057401	-13.32	0.000	-.08771	-.0652089
COPD	.0794734	.0072298	10.99	0.000	.0653031	.0936437
Covid	.4945386	.0050373	98.18	0.000	.4846655	.5044117
Pneumonia	.0539476	.0039335	13.71	0.000	.046238	.0616573
Respiratory_Failure	.3623239	.0097458	37.18	0.000	.3432222	.3814257
1.hiv	.033976	.0065779	5.17	0.000	.0210834	.0468686
1.ads	.014338	.0582273	0.25	0.805	-.0997871	.128463
1.ast	.0175278	.0038814	4.52	0.000	.0099203	.0251353
1.bmd	.0123541	.0095898	1.29	0.198	-.0064419	.0311501
1.bce	.0503876	.0196334	2.57	0.010	.0119063	.0888688
1.chf	.1457747	.0369374	3.95	0.000	.0733777	.2181717
1.cmy	-.0086604	.0072748	-1.19	0.234	-.022919	.0055982
1.cop	.0392913	.0091578	4.29	0.000	.021342	.0572406
1.crf	.0105811	.0126311	0.84	0.402	-.0141758	.035338
1.cad	-.0256057	.0115498	-2.22	0.027	-.0482432	-.0029681
1.chd	-.0167505	.2238792	-0.07	0.940	-.455552	.422051
1.db1	.0112707	.1736158	0.06	0.948	-.3290149	.3515564
1.dm1	.0357399	.0130805	2.73	0.006	.0101022	.0613777
1.dm2	-.0164746	.0050297	-3.28	0.001	-.0263328	-.0066164
1.dys	.0131518	.0104373	1.26	0.208	-.0073052	.0336089
1.epl	-.0099348	.0082095	-1.21	0.226	-.0260253	.0061558
1.glc	.010406	.0117473	0.89	0.376	-.0126186	.0334307
1.hae	.0090188	.1585143	0.06	0.955	-.301668	.3197056
1.hyl	-.0102742	.0051622	-1.99	0.047	-.0203921	-.0001562
1.hyp	-.027548	.0042422	-6.49	0.000	-.0358626	-.0192334
1.hyt	.1320122	.0623463	2.12	0.034	.0098139	.2542105
1.mss	.0142685	.062245	0.23	0.819	-.1077311	.1362681
1.par	.0111988	.0201056	0.56	0.578	-.028208	.0506056
1.rha	.0212556	.0098727	2.15	0.031	.0019052	.0406061
1.scz	-.0454414	.0273658	-1.66	0.097	-.0990782	.0081953
1.sle	.0051827	.0222638	0.23	0.816	-.0384542	.0488197
0.uci	0	(omitted)				
1.cancerid	.0590122	.009767	6.04	0.000	.039869	.0781554
1.obesityid	.125321	.0518933	2.41	0.016	.0236105	.2270314
1.tbid	-.01103	.0083288	-1.32	0.185	-.0273544	.0052944
Care_Level						
High Care	.5173515	.0054765	94.47	0.000	.5066176	.5280854
ICU	.8478982	.0084562	100.27	0.000	.8313241	.8644723
Net_work						
Hosp_Grp_2	.1642717	.0036989	44.41	0.000	.1570219	.1715214
Hosp_Grp_3	.1083296	.0036219	29.91	0.000	.1012307	.1154286
Hosp_Grp_4	.0491721	.0037449	13.13	0.000	.0418321	.0565121
_cons	8.957476	.0066106	1355.02	0.000	8.944519	8.970433

This model for total hospital benefits paid explains approximately 74% of the observed outcome and may be relied on (statistically significant with p-value <0.000).

The statistically significant risk factors are as follows:

- Admission reason
- Level of care
- Hospital network

Age was only statistically significant in those less than 35 years (p-value <0.001) while in those aged 35 or more it was not statistically significant.

Gender was not statistically different for the total hospital benefits paid (p-value = 0.083); therefore the total hospital benefits paid for males and females may be treated as similar.

Pre-existing conditions that were significant are: HIV; asthma; bronchiectasis; cardiac failure; chronic obstructive pulmonary disorder; coronary artery disease; diabetes mellitus type 1; diabetes mellitus type 2; hyperlipidaemia; hypertension; hypothyroidism; rheumatoid arthritis; cancer and obesity.

Similarly, for the purposes of constructing the final model, any patient having any one of these was classified as having a risk condition.

Therefore, the final model for total hospital benefits paid was as follows:

Source	SS	df	MS			
Model	36325.9145	33	1100.78529	Number of obs	=	84,229
Residual	12270.0571	84,195	.145733797	F(33, 84195)	=	7553.40
				Prob > F	=	0.0000
				R-squared	=	0.7475
				Adj R-squared	=	0.7474
Total	48595.9715	84,228	.576957443	Root MSE	=	.38175

T_TPD	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
LOS	.1266048	.0004784	264.67	0.000	.1256673	.1275424
AgeBand						
less than one year	-.0024138	.0077098	-0.31	0.754	-.0175249	.0126973
1-4 years	-.0332123	.0061482	-5.40	0.000	-.0452626	-.0211619
5-9 years	-.0441769	.0071297	-6.20	0.000	-.058151	-.0302027
10-14 years	-.0763333	.0083651	-9.13	0.000	-.0927288	-.0599378
15-19 years	-.1090567	.0087669	-12.44	0.000	-.1262397	-.0918737
20-24 years	-.0482659	.0100303	-4.81	0.000	-.0679253	-.0286065
25-29 years	-.0343606	.0099431	-3.46	0.001	-.053849	-.0148722
30-34 years	-.0117328	.0080379	-1.46	0.144	-.027487	.0040214
35-39 years	.0078491	.0072562	1.08	0.279	-.006373	.0220711
40-44 years	.0052995	.0069918	0.76	0.448	-.0084045	.0190035
45-49 years	-.0008531	.0065716	-0.13	0.897	-.0137334	.0120272
55-59 years	.0021168	.0064995	0.33	0.745	-.0106222	.0148557
60-64 years	.0041215	.0071373	0.58	0.564	-.0098675	.0181105
65-69 years	.0046581	.0079089	0.59	0.556	-.0108434	.0201595
70-74 years	.0101887	.0083123	1.23	0.220	-.0061034	.0264807
75-79 years	.0140705	.008868	1.59	0.113	-.0033108	.0314517
80-84 years	-.0008951	.0098414	-0.09	0.928	-.0201841	.0183939
85 years+	.007557	.0101113	0.75	0.455	-.012261	.027375
Sex						
M	.003414	.0027473	1.24	0.214	-.0019707	.0087986
Rk_CDL						
No	-.0109144	.0074768	-1.46	0.144	-.0255687	.00374
Care_Level						
High Care	.5160421	.0054207	95.20	0.000	.5054175	.5266667
ICU	.8437754	.0083752	100.75	0.000	.82736	.8601907
Admin_Reason						
Bronchitis	-.0808123	.0055511	-14.56	0.000	-.0916923	-.0699323
COPD	.0770389	.0067117	11.48	0.000	.0638841	.0901937
Covid	.4306005	.0051327	83.89	0.000	.4205404	.4406606
Pneumonia	.0363783	.0036952	9.84	0.000	.0291357	.0436209
Respiratory_Failure	.2798548	.0096264	29.07	0.000	.2609872	.2987223
Net_work						
Hosp_Grp_2	.1662168	.003649	45.55	0.000	.1590647	.1733689
Hosp_Grp_3	.1053816	.0035691	29.53	0.000	.0983861	.112377
Hosp_Grp_4	.0512016	.0036959	13.85	0.000	.0439577	.0584456
Year						
2020	.0715301	.0031168	22.95	0.000	.0654211	.077639
CDL_Count	-.01571	.0073826	-2.13	0.033	-.0301798	-.0012402
_cons	8.905417	.0096183	925.89	0.000	8.886565	8.924268

This modified model for total benefits paid explains 75% of the observed results (R^2=75%) and is statistically significant (p-value < 0.000). The only differences with original model are age and that the risk pre-existing medical condition indicator which is not statistically significant (p-value = 0.14). The total benefits paid for patients under the age one year are not statistically different from the total benefits paid for patients above 30 years of ages (p-values = 0.754).

9.4. ANNEXURE 4: Goodness of fit of model

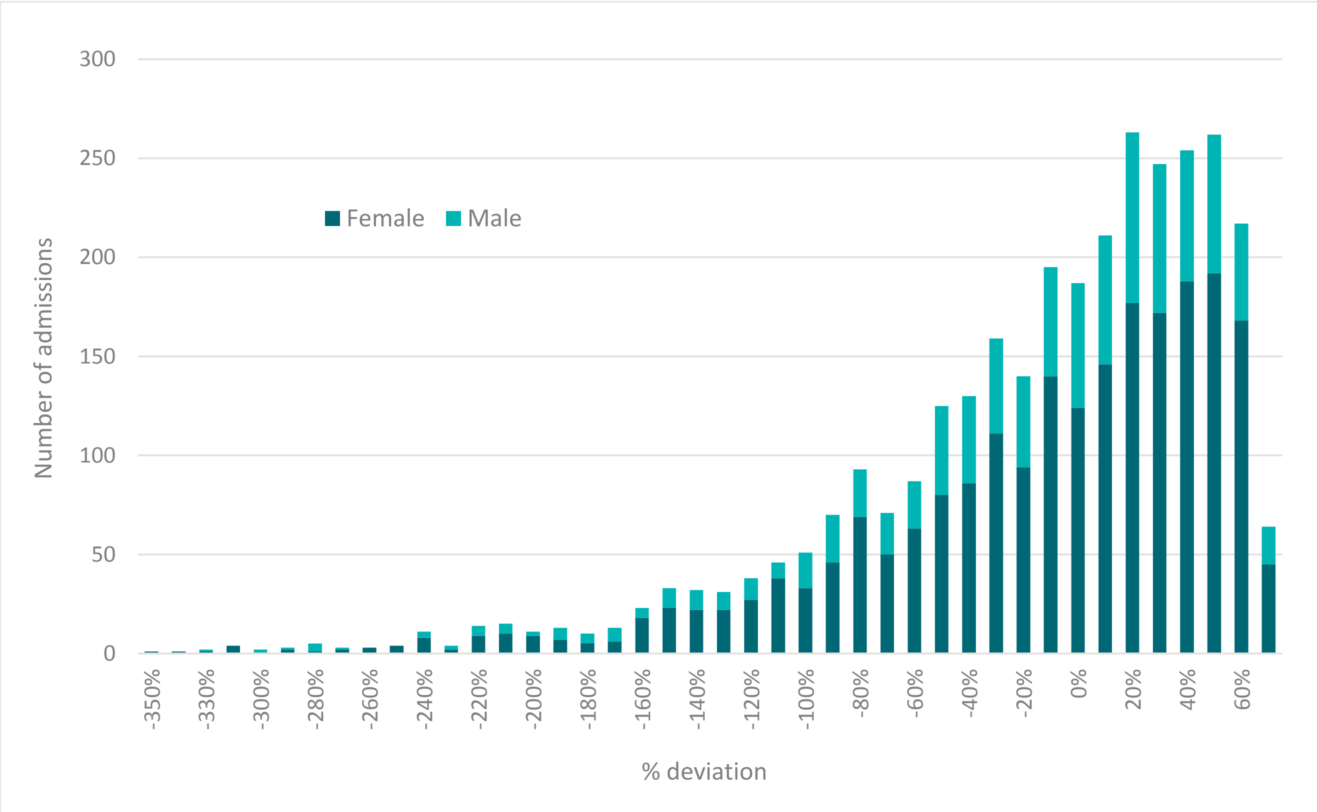
ANNEXURE 4.1. Total benefits paid

The predicted total benefits paid were compared to the observed total benefit paid and the following graph shows the percentage differences. The results shown are filtered as follows:

- Patients with at least one risk condition
- General ward admissions only
- COVID-19 admissions only

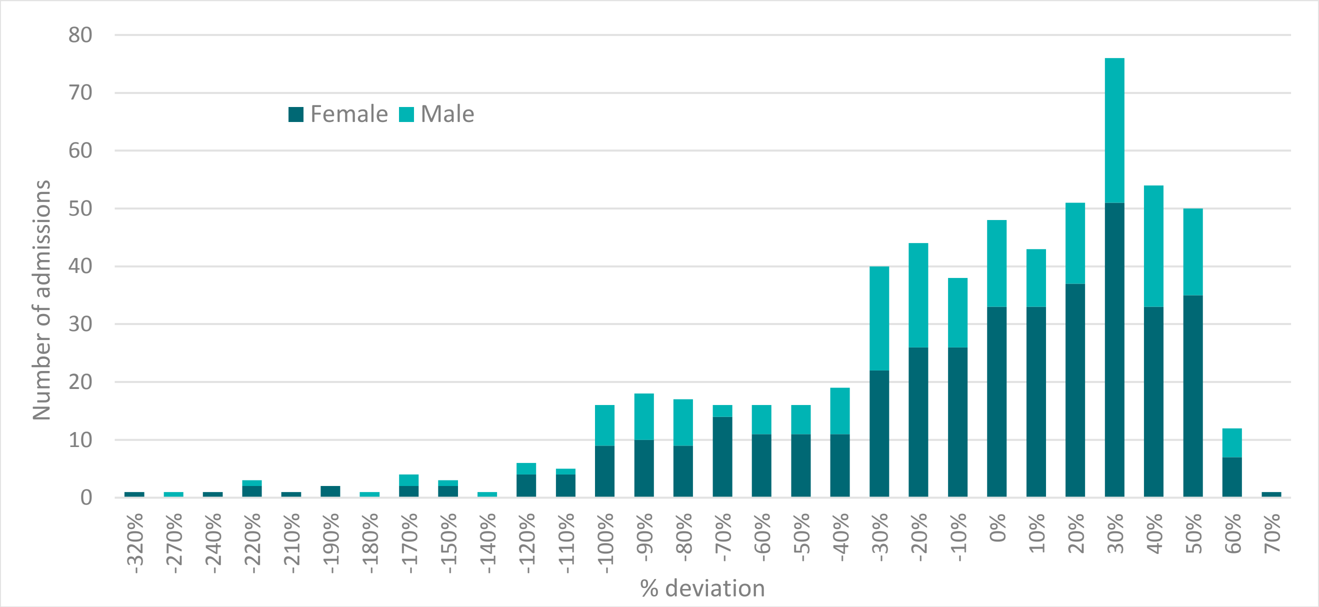
Asthma

Figure 21: Asthma model goodness of fit; total benefits paid



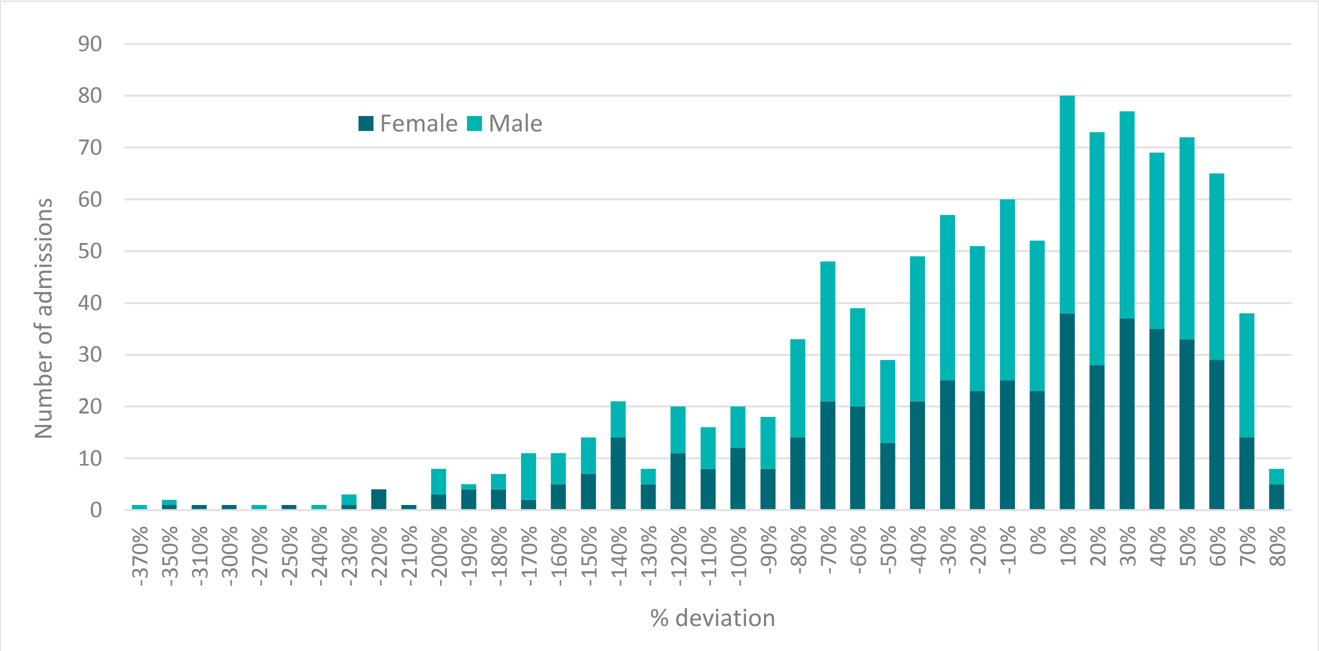
Bronchitis

Figure 22: Bronchitis model goodness of fit; total benefits paid



COPD

Figure 23: COPD model goodness of fit; total benefits paid

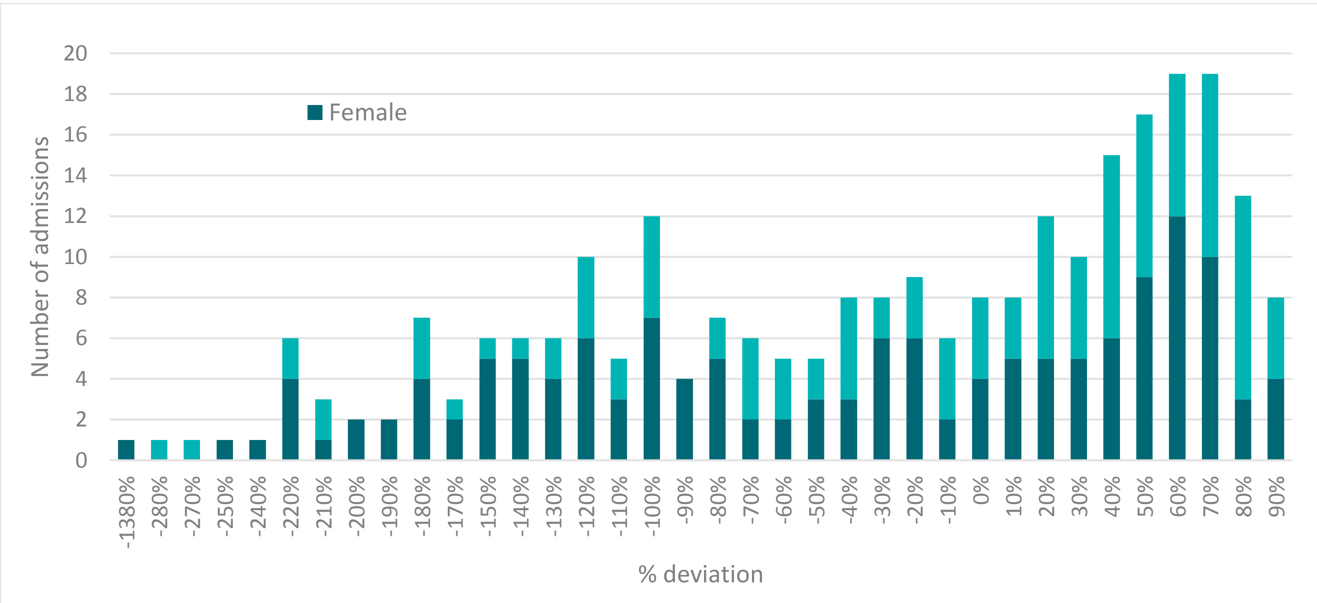


Pneumonia

Figure 24: Pneumonia model goodness of fit; total benefits paid

Respiratory failure

Figure 25: Respiratory failure model goodness of fit; total benefits paid



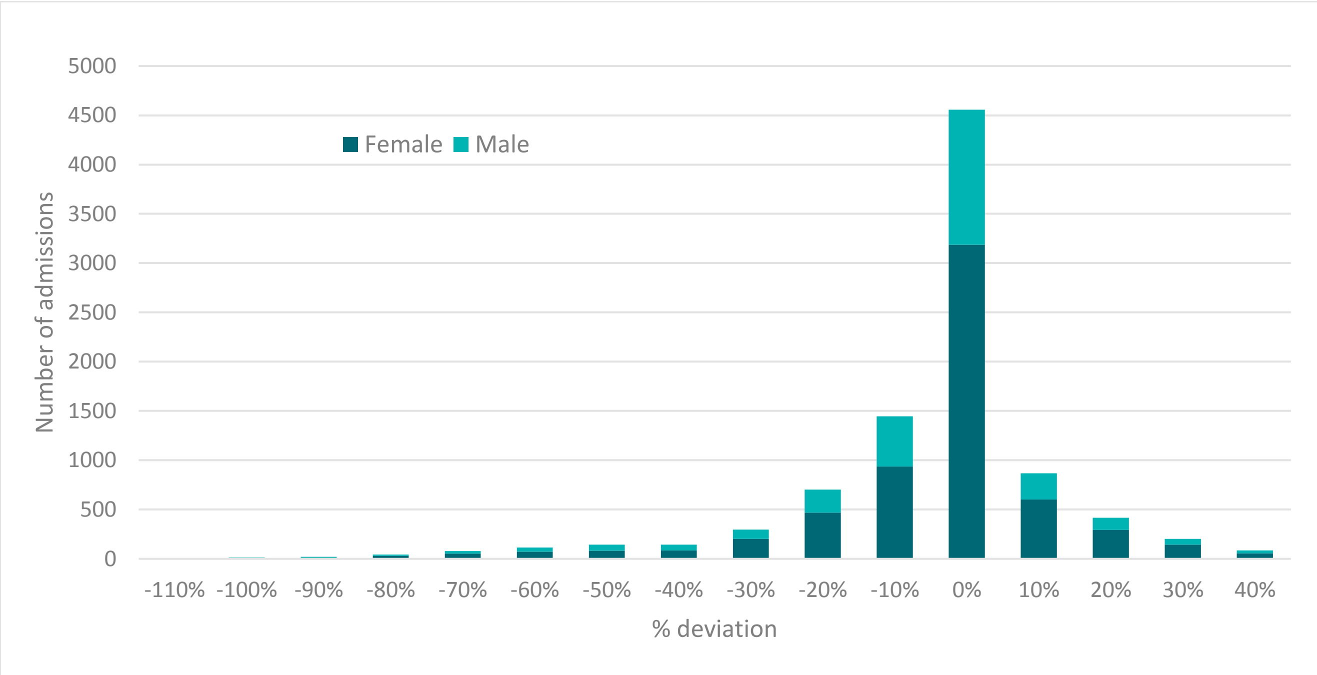
Annexure 4.2. Average daily cost

The predicted average daily costs were compared to the observed average daily costs and the following shows the percentage differences. The results shown are filtered as follows:

- Patients with at least one risk condition
- General ward admissions only
- COVID-19 admissions only

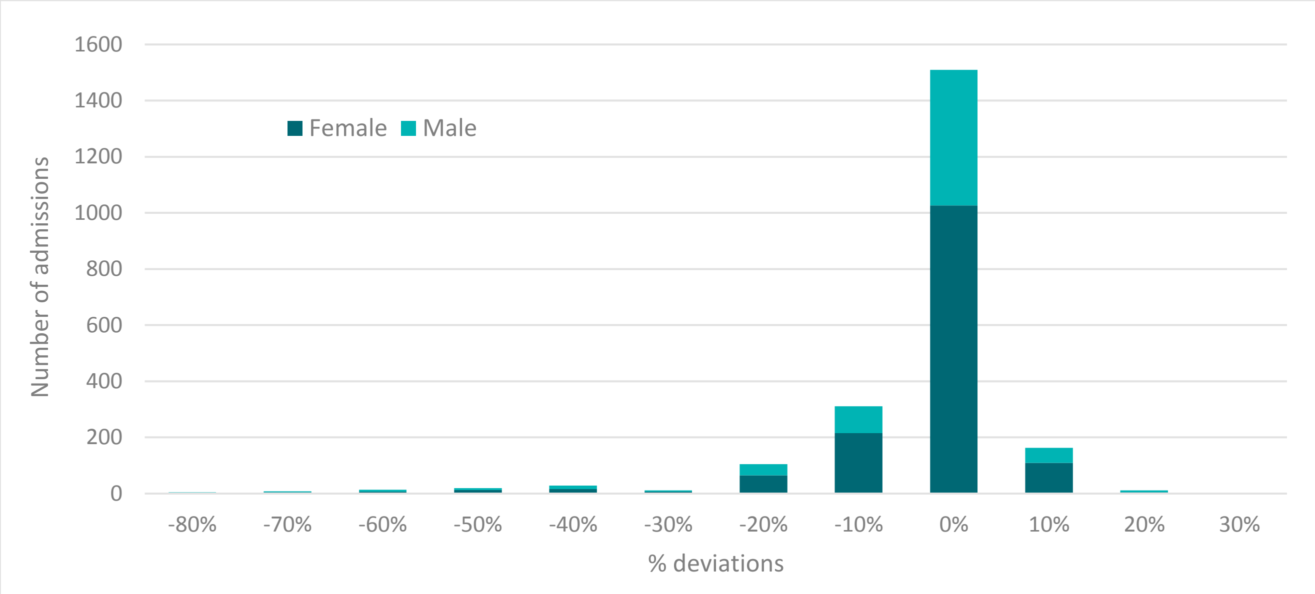
Asthma

Figure 26: Asthma model goodness of fit; average daily rate



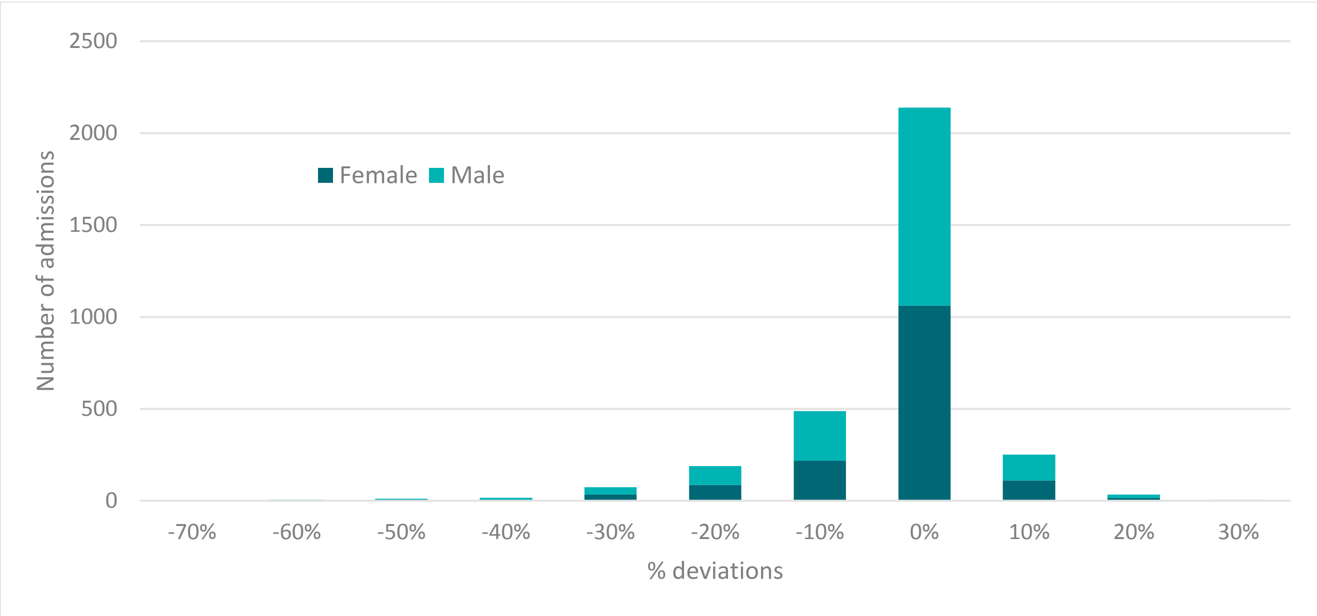
Bronchitis

Figure 27: Bronchitis model goodness of fit; average daily rate



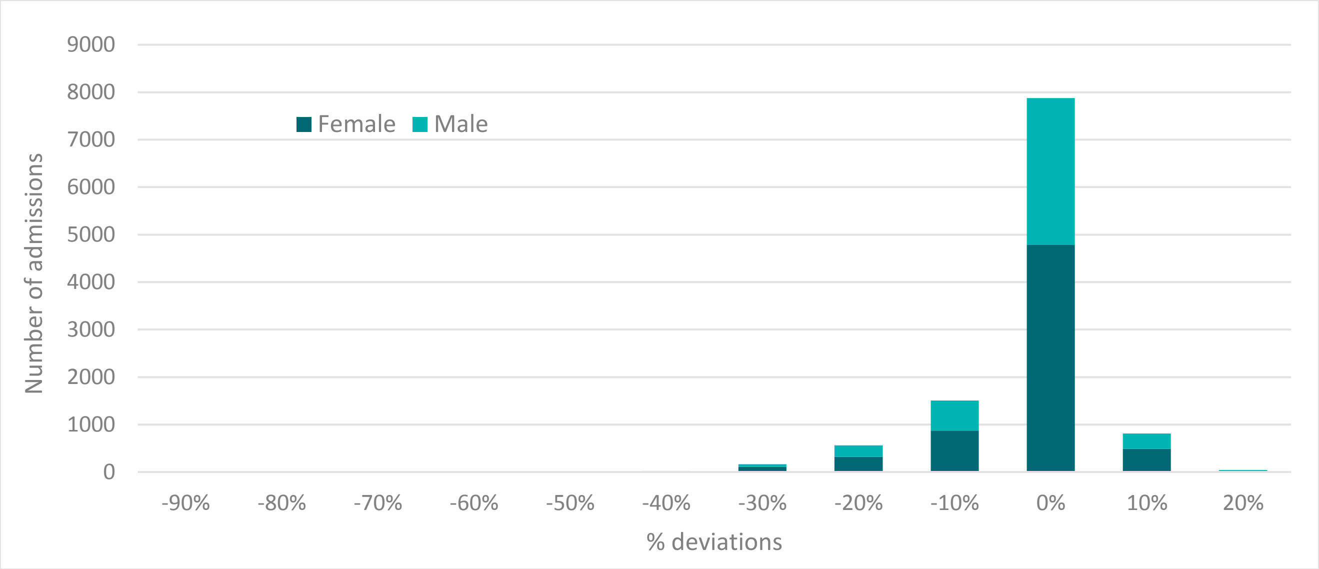
COPD

Figure 28: COPD model goodness of fit; average daily rate



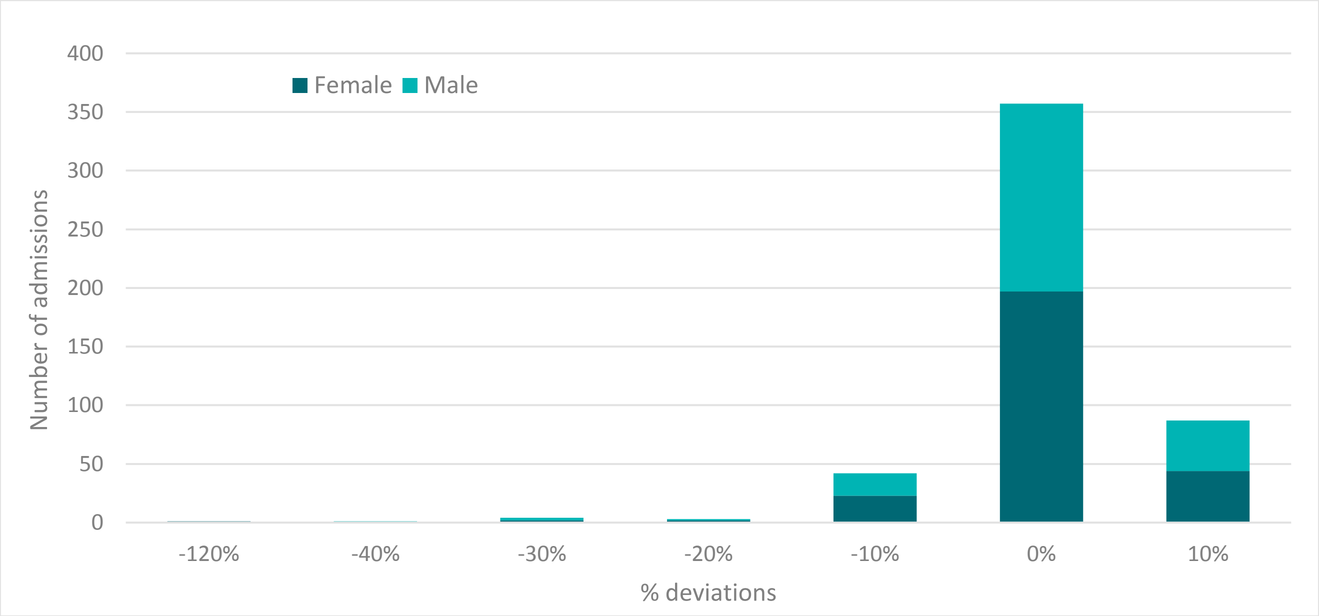
Pneumonia

Figure 29: Pneumonia model goodness of fit; average daily rate



Respiratory failure

Figure 30: Respiratory failure model goodness of fit; average daily rate



CONTACT DETAILS

Charlton Murove

Head: Research

T: +27 11 537 0256

M: +27 76 812 2780

CharltonM@bhfglobal.com

Board of Healthcare Funders

http://www.bhfglobal.com

South Tower, 160 Jan Smuts Ave,
Rosebank, 2196

B·H·F
BOARD OF HEALTHCARE FUNDERS