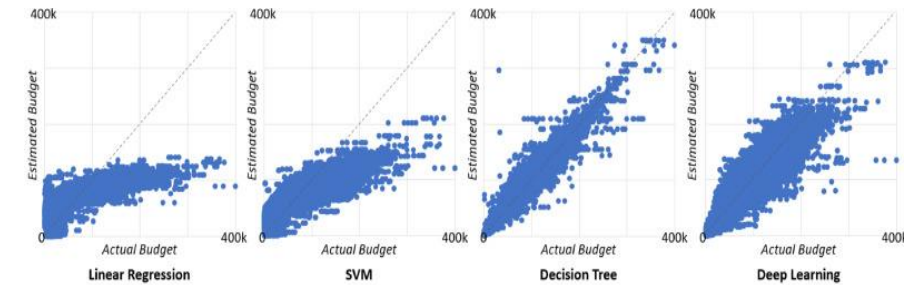


# Health Insurance Claim Prediction Using Artificial Neural Networks (ANN)

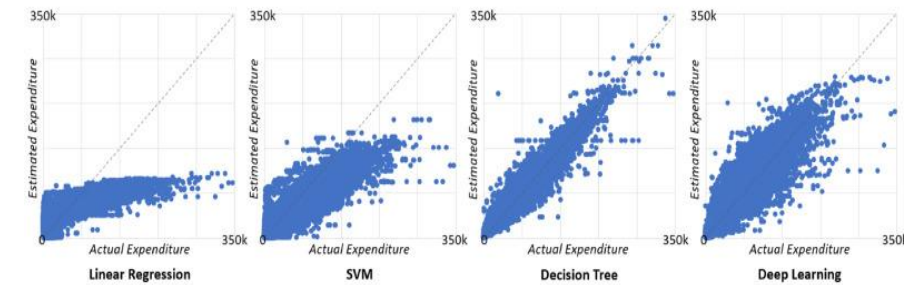
- **Goal:** This study presents the development of artificial neural network model that is appropriate for predicting the anticipated annual medical claims.
- **Dataset:** The data consists of the 20 years (1999-2018) of raw claims data provided by BSP LIFE (Fiji) Ltd. The raw data was summarized yearly to explore the trends in the yearly claimed amount paid.
- **Methodology:** Raw data was divided into training and test data in the ratio of 70% and 30%. Claim records from 1999 to 2011 were used for training and records from 2012 to 2018 were used for testing. The research steps involved taking the raw data, performing an auto-correlation and then initiating the data normalization followed by training the test data. Finally, network design was checked for network accuracy and the stopping criteria. MATLAB software was used to plot the auto-correlation of the series, which highlighted the positive correlation in 1, 2 and 3 lag variables. The Lags were used to identify the number of input to the network i.e., the number of years of claims amount needed as the input. Epochs and MAPE was used to stop the training and testing process. The neural network models were tested with a number of different Epochs where MAPE is observed. For each Epoch size, MAPE was observed and eventually reached a point where the lowest MAPE was observed. Root Mean Squared Error on the training data was also used as the stopping criteria.
- **Results:** The comparison of the result with Human prediction clearly shows that ANN outperforms in prediction. It reduced the error by 11.5%. It can be concluded that ANN can be used for medical claims prediction resulting in strong forecasting.
- **Article Link:** <https://www.igi-global.com/article/health-insurance-claim-prediction-using-artificial-neural-networks/257242>

# Learning from machines to close the gap between funding and expenditure in the Australian National Disability Insurance Scheme

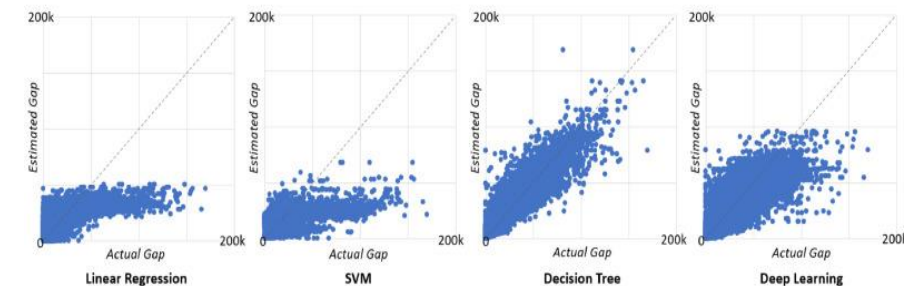
- **Goal:** Using machine learning techniques to estimate budget and close lose funding gap for the Australian National Disability Insurance Scheme (NDIS).
- **Dataset:** Total 8 quarters of NDIS data is used from Q3 2019 to Q2 2021.
- **Methodology:** The 8 Quarters NDIS data is analyzed to show the large gap to between the allocation and utilization. To address this gap, four machine learning models were trained and tested using the NDIS data for the estimation of NDIS budget, expenditure and gap by MAE and RMSE evaluation. After that the best performing model was selected and used to simulate a machine learning-based approach to close the gap over time. To understand the advantages of using this approach, it was then compared with human learning approach that is currently employed in the NDIS. Lastly, the decision tree hierarchy and value analysis was used to explore the significant factors used by machine learning models in the estimations.
- **Results:** The gap between the allocated budget and realized expenditure in NDIS can be closed faster and at a reduced cost using an appropriate machine learning model compared to the current manual processes.
- **Article Link:**  
<https://www.sciencedirect.com/science/article/pii/S2667096822000209>



(a) Scatter plots showing the estimated budget by the four machine learning models compared to the actual budget.



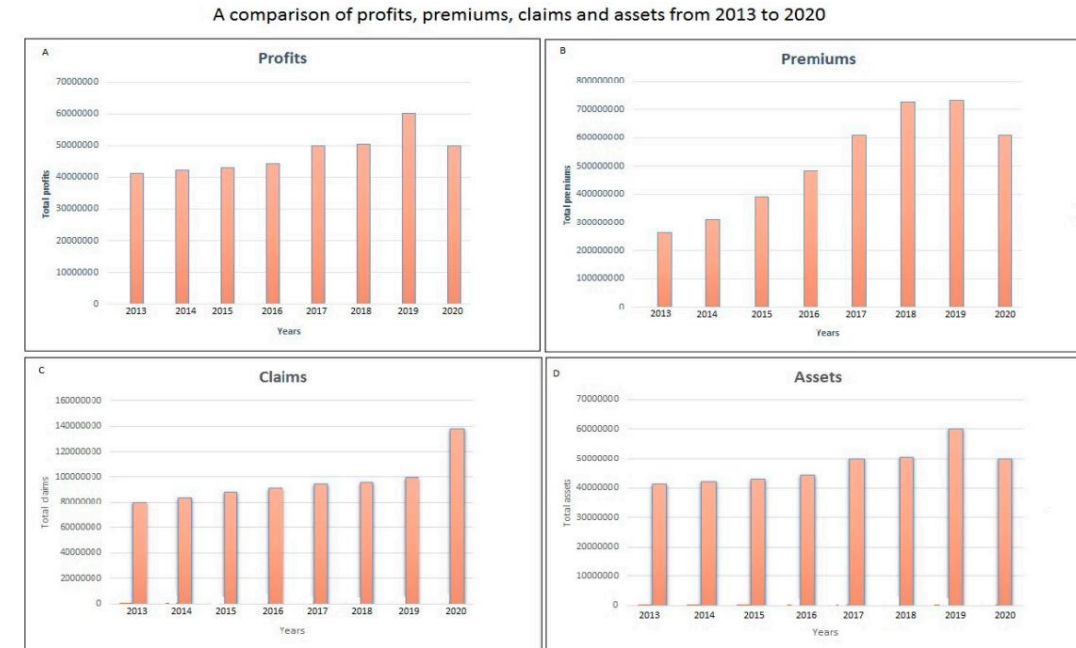
(b) Scatter plots showing the estimated expenditure by the four machine learning models compared to the actual expenditure.



(c) Scatter plots showing the estimated gap by the four machine learning models compared to the actual gap.

# The Impact of COVID-19 on the Insurance Industry

- **Goal:** To estimate the impact of COVID-19 on the insurance industry in Ghana & to discuss insurer expectations and present solutions.
- **Dataset:** Insurance data was collected as far back as 2013 to compare trends of operations. Data for this study include operations data and documents from insurance companies, Ghana Health Service (GHS), World Health Organization (WHO), Ghana Statistical Service Department (GSS), Ghana National Insurance Commission (NIC), Ghana Labour Commission (LC) and the Ghana National Health Insurance Authority (NHIA). From the insurance companies, the data consist of total premiums, claims, profits, assets and liabilities, market share and business investments.
- **Methodology:** Three different types of questionnaires to three types of responders were designed and Qualitative and quantitative interviews were administered to capture the level of risk the pandemic posed on insurance companies; responses of insurers; actions and possible solutions; expectations of insurers and their preparation.
- **Results:** A comparison of 2013 to 2020 indicates reducing profits, premiums, and assets but increasing claims. However, comparison and forecast predicts a normalization of economic indicators from January 2021.
- **Article Link :** <https://www.mendeley.com/catalogue/f887344e-9549-3f80-966c-c308e0c11751/>



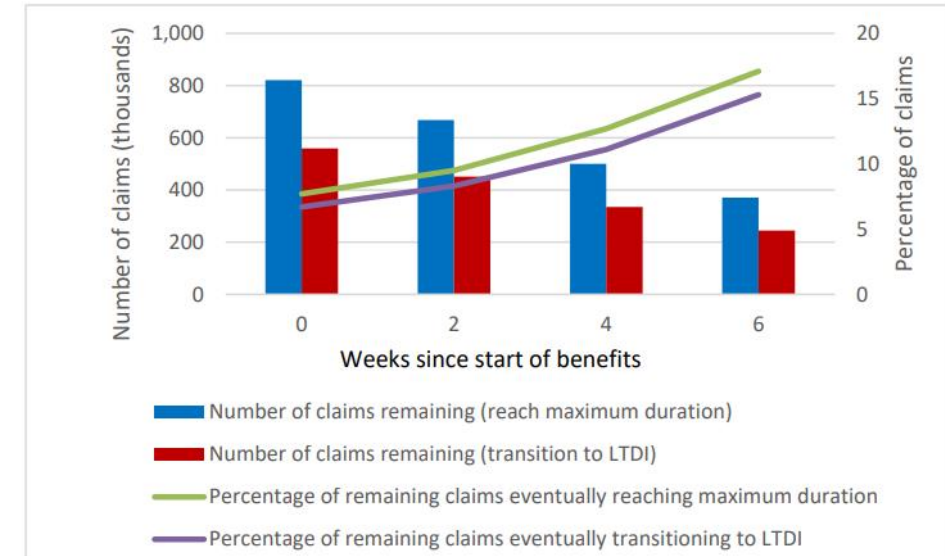
**Figure 1.** General quarterly (March to June) outlook of the insurance industry: (A) trend of profits, (B) trend of premiums, (C) trend of claims and (D) trend of assets.

# Transgender-Related Insurance Denials as Barriers to Transgender Healthcare: Differences in Experience by Insurance Type

- **Goal:** This study examine the association between multiple forms of TNB insurance–based denial and type of insurance. It investigates the association between transgender and non-binary individuals' experiences of different forms of transgender-related insurance denials and insurance type.
- **Dataset:** This paper utilizes secondary USTS data which has received exempt approval from the University of Michigan IRB. The 2015 United States Transgender Survey was conducted by the National Center for Transgender Equality to ascertain US transgender and non-binary experiences across multiple life experiences, and in partnership with over 400 organizations with over 27,715 participants.
- **Methodology:** Multivariate logistic regressions were conducted, and adjusted risk ratios were calculated, to analyze the likelihood of experiencing eight different forms of denials by insurance type, including private, Medicare, Medicaid, and military-related, and having multiple types of insurance coverage. Multivariate logistic regression results are presented for all eight denial variables. All eight models showed overall statistical significance (all  $p < 0.001$ ), and Hosmer-Lemeshow goodness of fit tests were insignificant for all models, indicating good fit and that actual and predicted frequencies were not statistically different (all  $p > 0.05$ )..
- **Results:** : Models revealed significant relationships between transgender-related insurance denials and insurance type for 11,320 transgender and non-binary adults. Compared with those with private insurance, Medicaid coverage was associated with an increased likelihood of experiencing denials for hormone therapy (adjusted risk ratio (ARR) = 1.22; CI = 1.05–1.42;  $p = 0.02$ ); having no in-network surgery providers was associated with Medicare (ARR = 1.84; CI = 1.29–2.62;  $p = 0.009$ ) or Medicaid (ARR = 1.54; CI = 1.20–1.98;  $p = 0.003$ ); and military based insurance was associated with transition-related surgery denials (ARR = 1.53; CI = 1.36–1.72;  $p < 0.001$ ).
- **Article Link:** <https://www.mendeley.com/catalogue/8140157b-a1ae-3eb0-98f1-818ba6a44f36/#author%20supplied%20keywords-title>

# Use of Predictive Analytics in Short-Term Disability Claim

- **Goal:** Using Predictive Analytics for Early Identification of Short-Term Disability Claimants Who Exhaust Their Benefits
- **Dataset:** Insurance Claim data from Integrated Benefits Institute (IBI) Health and Productivity Benchmarking Data from 2011 through 2015, including 820,751 close STD claims from 8,587 small, medium, and large businesses associated with 9 disability insurance carriers and third-party leave administrators.
- **Methodology:** Several predictive models were fit to the data, including logistic regression, regularized logistic regression (using an elastic net), and random forests. Samples were randomly divided into training and testing sets, and selected a model based on the area under the receiver operating characteristic curve. Individuals were flagged as having a high probability of exhausting their benefits using a predicted probability threshold, which was chosen to balance the tradeoff between sensitivity and specificity.
- **Results:** In the logistic regression, the factors most strongly associated with exhaustion of STD benefits are age, diagnosis, and employer industry. Waiting to allow some claims to resolve without intervention improves the efficiency of targeting efforts..
- **Article Link:**  
[https://appam.confex.com/data/extendedabstract/appam/2018/Paper\\_26234\\_extendedabstract\\_1780\\_0.pdf](https://appam.confex.com/data/extendedabstract/appam/2018/Paper_26234_extendedabstract_1780_0.pdf)



Predicting which claims exhaust benefits



# Infants without health insurance

- **Goal:** It is a preliminary exploration of how the uneven expansion of Medicaid affects racial/ethnic health insurance coverage across the rural/urban spectrum in the United States.
- **Dataset:** The specific data used in the analysis were the 2011–2015 five-year data downloaded from the Integrated Public Use Microdata Series (IPUMS) U.S. database in order to identify infants and their mother's specific data extracted from ACS five-year estimates for 2011–2015 microdata files that have a sample size of 15,637,457.
- **Methodology:** Used supervised machine learning models' Logistic regression & Multivariate logistic regression to predict probabilities of each interaction. Generating a multivariate logistic regression that estimated the likelihood of not having health insurance and controlled for maternal demographics and other household characteristics. The logistic regression was appropriate for our purposes because the probability of an infant being born to a mother with or without health insurance is dichotomous. The predicted probabilities of each interaction were then graphed for improved interpretation because even though a significant interaction variable in a logistic regression model indicates that the effect of a variable is not the same for all values of the interacting variable, the reported values do not provide clear information about the nature of the interaction. Graphing the interaction provides a more complete understanding of the nature of the interaction.
- **Results:** The overall results indicated that approximately 17.4% (572,277 weighted households) of all infants are born into households with no form of health insurance. A nearly significant ( $p < .06$ ) discrepancy emerged between rural and urban infants, in which approximately 19.9% (70,224 weighted households) of rural infant households had no health insurance compared to 16.8% (502,053 weighted households) of urban infant households
- **Article Link:** <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0222387>

# Predicting Disability among Community-Dwelling Medicare Beneficiaries Using Claims-Based Indicators

- **Goal:** To assess the feasibility of using existing claims-based algorithms to identify community-dwelling Medicare beneficiaries with disability based solely on the conditions for which they are being treated, and improving on these algorithms by combining them in predictive models
- **Dataset:** Data on 12,415 community-dwelling fee-for-service Medicare beneficiaries (10,057  $\geq$  age 65 and 2,357 ages 18-64) who first responded to the Medicare Current Beneficiary Survey (MCBS) in 2003–2006.
- **Methodology:** A logistic regression model was estimated for each age where disability indicator was used as a dependent variable while the six claim-based disability indicators were main predictors. For each estimated model, ROC curve was constructed that showed all possible sensitivity–specificity combinations for that model. ROC analysis evaluates the predictive performance of a model by capturing the model's trade-off between sensitivity and specificity.
- **Results:** To identify people with disability in claim data predictive models provide an improved tool when multiple claim-based indicators are considered. For age 65 or older sensitivity–specificity trade-off for the models is considerably better than for those ages 18-64.
- **Article Link:** <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4722209/>

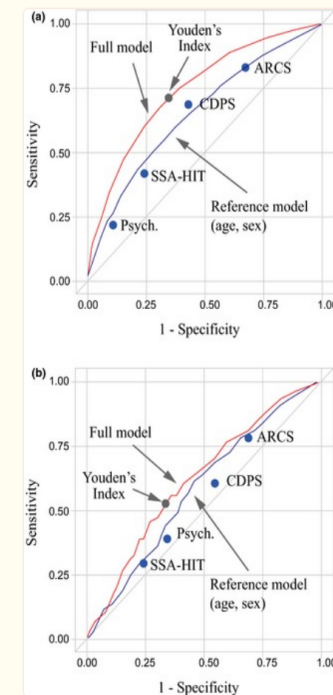


Figure 1

ROC Curve versus Claims-Based Disability Flags, Limitation in at Least One ADL: (a) Age 65 or older; (b) Age 18-64

# Health and disability insurance

- **Goal:** This paper is based on an extract of variables which include whether a person receives disability insurance or not; basic demographic characteristics, and a broad set of health variables. Additionally, this paper investigates the causes for this variation about Why are so many more individuals aged 50–64 receiving DI benefits in Denmark, Sweden and the Netherlands than in e.g., France or Germany? Why so many fewer individuals in Greece than in Poland?
- **Dataset:** SHARE wave 3 covers 13115 such individuals. ELSA and HRS contribute 6732 and 4270 individuals, respectively, to the joint sample, thus consisting of up to 24117 individuals. This paper is to shed light on this trade-off by using three waves of the Survey of Health, Ageing and Retirement in Europe (SHARE), together with data from its sister surveys in England (the English Longitudinal Study on Ageing, ELSA) and the US (the Health and Retirement Study, HRS).
- **Methodology:** Using joint descriptive statistics, the calibrated weights have been re-normalized to give each country equal weight. Three regressions were used to analyze: a simple linear model for the probability to become enrolled into disability, a probit specification, and a logit specification. A first finding is the similarity among the three specifications. A second observation is the large unexplained variation. This is in line with the findings of OECD (2003) where only little correlation between “medical disability status” and “disability enrolment status” was found.
- **Results:** This paper does not provide significant correlation with objective health measures. It shows a weak correlation with self-rated health may be influenced by “justification bias”, . It concludes that it is not health but the country-specific design of early retirement and Labour market institutions, and especially disability insurance rules, which explain the observed cross-country variation in the receipt of disability benefits.
- **Article Link:** <https://labourmarketresearch.springeropen.com/articles/10.1007/s12651-011-0058-6>