

# DISCOVERY HEALTH: LEVERAGING DATA SCIENCE AND MACHINE LEARNING TO GAIN INSIGHTS AND SAVE LIVES

**IMPACT**

- 48% reduction in mortality rate for patients who received a pulse oximeter
- 20,000 additional members identified as high risk, compared with the standard classification.
- 900 data points combined to make more accurate predictions that helped save lives

Discovery Health is a South African health insurance company whose purpose and ambition are achieved through a pioneering business model that incentivizes people to be healthier and enhances and protects their lives. With a data-driven approach, its shared value insurance model delivers better well-being and value for clients, superior actuarial dynamics for the insurer, and a healthier society.

Today, Discovery Health is the leading medical scheme administrator in South Africa, providing administration and managed care services to over 3.5 million beneficiaries, including 18 restricted medical schemes on behalf of leading corporate clients, as well as Discovery Health Medical Scheme, South Africa's largest open medical scheme. Discovery Vitality is a dynamic and science-based behavior change program that combines data analytics with rewards and incentives to help its customers make healthier choices. Data collected includes fitness device data, including vital statistics such as heart rate, sleeping patterns, and fitness activity. The exponential growth of the company's data and data sources over the last few years led it to seek a partner who could help improve the analysis of its large quantities of data.

"The partnership with Cloudera was a natural one for us as we saw more enterprises investing in data science and we wanted to embrace our big data journey," said Shirley Collie, Chief Health Analytics Actuary, Discovery Health. The magnitude of how Discovery Health could use its data for good was realized when the COVID-19 pandemic hit. "The COVID-19 pandemic presented a 'once in a century' challenge. We wanted to do all we could to help our customers through it."

Discovery Health has been a key stakeholder in South Africa's response to COVID-19.

### The COVID-19 Challenge

When the pandemic started impacting South African's day-to-day lives, the Discovery Health team identified a decreased rate of people seeking healthcare. The company experienced a drop in phone calls coming through and fewer claims. Compared to the same period in the previous year, there were 30% fewer general practitioner (GP) consultations among medical scheme members who lived with chronic illness.

At the start of the pandemic prior to high COVID-19 case-loads in South Africa, the team came across a New York Times article written by a doctor working in an A&E department during the peak of coronavirus. The piece explained how a simple device, such as a pulse oximeter, could decrease mortality rates in high-risk patients. Discovery Health realized they had substantial data at their fingertips, which could help to identify clients who might be at a higher risk of hospitalization, and allow them to take proactive, preventative measures like this and help to improve their health.

# 48%

Reduction in mortality rate for patients who received a pulse oximeter

“Working with Cloudera’s data platform has enabled our machine learning models to prioritize those individuals who are most at risk of ending up in the hospital, helping preserve the healthcare ecosystem and appropriately use the resources that were becoming so scarce in the face of the virus”

– Shirley Collie, Chief Health Analytics Actuary, Discovery Health

## Using data to save lives

This spurred the team into action and the result was The Pulse Oximeter Outreach initiative, deployed on Cloudera Data Science Workbench (CDSW), in order to put their theory into practice. “Our main objective was to ensure we were using our data in the best possible way, in order to be a force for social good,” said Collie.

Discovery Health deployed their Patient-X© pipeline machine learning model to identify individuals that would be most likely to be at-risk of hospital admission, should they contract the virus. CDSW allowed the team to identify these patients and proactively target them with preemptive communications containing advice, in order to try and prevent the need for hospital admission. It also enabled them to prioritize patients that should receive a pulse oximeter device, which would monitor oxygen in real-time and help to flag issues earlier. This was particularly useful once the number of infected patients started to rise. Discovery Health identified an incredible 48% reduction in mortality risk for patients who received a pulse oximeter.

“Working with Cloudera’s data platform has enabled our machine learning models to prioritize those individuals who are most at risk of ending up in the hospital, helping preserve the healthcare ecosystem and appropriately use the resources that were becoming so scarce in the face of the virus,” commented Collie.

As a single integrated platform connecting the entire data lifecycle, Cloudera offered Discovery Health the chance to combine 900 different data points consisting of demographic, lifestyle, and clinical features, to make better informed predictions. Furthermore, Cloudera’s consistent governance around POPIA, HIPAA and GDPR compliance allowed Discovery Health to make sure its clients’ privacy remained protected, whilst being able to use sensitive, personal customer data to help take life-changing measures.

The secure use of these combined sets of data, from previous medical history to logged exercise, helped determine new factors that could contribute to a heightened risk level. These were additional to the commonly known, generic priorities for COVID-19, such as the elderly, and those with known, underlying health conditions.

For example, the team was able to identify that a younger person on the system with diabetes and a history of health issues should be prioritised over a 70-year-old man that regularly exercised and was in good health. At-risk patients were then contacted over the phone, and if necessary, received pulse oximeters as a precaution, or provided advice on lifestyle changes that may reduce their risk factors. Lower risk individuals were sent an email notification with the best precautions they could take. Thanks to the combined use of data sets, the Patient-X© pipeline was proven remarkably accurate when backtested against actual admissions of positive COVID-19 patients. Furthermore, 20,000 additional members were identified as high risk using the Patient-X© pipeline, compared with the standard clinical rules-based classification.

**About Cloudera**

At Cloudera, we believe that data can make what is impossible today, possible tomorrow. We empower people to transform complex data into clear and actionable insights. Cloudera delivers an enterprise data cloud for any data, anywhere, from the Edge to AI. Powered by the relentless innovation of the open source community, Cloudera advances digital transformation for the world's largest enterprises.  
Learn more at [cloudera.com](https://cloudera.com)

**The result: preemptive care to save lives in a pandemic**

As a second phase of Discovery Health's COVID-19 initiative, the company was able to use the insights gathered with the Patient-X© pipeline to create the COVID Resilience Index©. Available on Discovery Health's website, the tool uses analytics to calculate each member's risk factors, by analysing lifestyle data the company has from its Vitality members, such as logged exercise, nutrition patterns, mental health insights and healthy lifestyle indicators. The Index helps inform members on the best ways to protect themselves, backed up by real data.

With the COVID Resilience Index© Discovery Health was able to get further insights on the significant impact that exercise had on morbidity and mortality rates of COVID-19. The tool showed, for example, that a 65-year-old male, with no chronic conditions, who exercises for half an hour at least four times a week, has the same mortality risk from COVID-19 as a 45-year-old who exercises once a week.

Once the urgency of identifying, monitoring and contacting high risk patients dials down, the Discovery Health team will be able to easily shift its mindset towards prediction. Using the same data sets already established on the Cloudera platform, Discovery Health is prepared to minimise admission and mortality rates for any future outbreaks. This experience enables the company to expand their models further for different healthcare use cases. For example, it can use the models to help detect the likelihood of someone developing mental health challenges, such as depression, based on health-related data sets.

With the Cloudera data platform as the backbone of its data strategy, Discovery Health will be able to continue to focus on finding explanations for past events. It will also be able to ask, and answer, questions that can help shape the future – for example, which clients could develop diabetes and how that could be changed? Ultimately, through the use of data, Discovery Health can pre-empt care instead of remediating and gain the ability to help to save more lives.