Response to 'A critical analysis of Discovery Health's claims-based risk adjustment of mortality rates in South African private sector hospitals' by Rodseth *et al.*

To the Editor: We write to you regarding 'A critical analysis of Discovery Health's claims-based risk adjustment of mortality rates in South African private sector hospitals' by Rodseth *et al*. published in the *SAMJ* January 2023 edition, vol. 113 issue 1.[1] The article is authored by employees and representatives of Netcare Limited, and critiques a previous article published by Discovery Health in the *SAMJ* in 2019[2] describing claims-based risk adjustment of mortality rates in South African (SA) private sector hospitals.

We welcome all additional scientific research and debate that will lead to improved measurement and quality of care across the health system, and further embrace debate around our article to achieve consensus on this important topic. In this light, we felt it important to address a number of the points presented in the Netcare article. The following assertions made by Rodseth *et al*.[1] are of concern and should be addressed in the context of the extensive efforts to ensure quality of care in all hospitals across SA.

Measurement of outcomes from administrative data sources

Administrative data are widely used and proven to accurately identify and measure variation in health outcomes.[3] Administrative data have the additional advantages of offering a larger quantum of data and cases, and ensuring consistency of data across patients.[4] Several studies globally[5-9] have demonstrated a strong correlation in clinical outcome and process indicators across administrative and clinical databases, validating the use of administrative data in measuring quality of care. In the absence of widespread implementation and adoption of electronic health records in private hospitals across SA, the use of available administrative data to determine variation in the quality of the healthcare services provided is critically important. Rodseth *et al*.'s assertion that the measurement of outcomes from administrative data sources is incomplete fails to acknowledge the value of comprehensive administrative data against the backdrop of largely incomplete electronic health record-keeping. It is pleasing and welcomed that hospital groups are now investing in digitisation, which will undoubtedly improve accessibility of clinical data for future studies relating to both quality and efficiency of care.

Selection of factors included in the analytical models

Rodseth *et al*. suggest that selection of factors used in the models is neither clinically validated nor intuitively obvious. This overlooks the fact that the Discovery Health methodology selected factors based on both clinical relevance and uplift to the model prediction accuracy.[2] The factors included in each of the models are well established clinically relevant risk factors, which are proven to have a strong influence on the outcomes. These include demographics (age, sex), acute risk predictors based on admission World Health Organization ICD10 coding and chronic risk predictors, consistent with well-established methodologies.[10]

Risk adjustment including scoring that predicts mortality outcomes

The Discovery Health risk adjustment models include the Truven disease staging grouper (DSG),[11] a tool validated and designed to predict the risk of mortality, owing to its importance and power in predictive models for mortality and morbidity. This DSG is similar to the alternative scoring systems put forward by Rodseth *et al*. The DSG is a subcomponent of the medical episode grouper (MEG), which is the proprietary episode grouping methodology of Merative (previously Truven Health Analytics, an IBM Watson (USA) company). Today, >190 health plans, employers and state Medicaid agencies use MEG to compare and contrast medical and surgical options and costs in the treatment of diseases and medical conditions.

Model redundancy and multi-collinearity

When the primary purpose of a model is prediction rather than inference, multi-collinearity and redundancy are not significant statistical issues. In fact, some studies suggest that including highly correlated variables in a prediction model improves its accuracy, and that machine learning and data-driven approaches may be less sensitive to these issues.[12-15]

Technical information and sensitivity analysis

The Discovery Health article[2] is an explanation and examination of variation in standardised mortality rates at hospital system level, and does not attempt to provide an exhaustive technical description of the methodology. Discovery Health has made significant efforts over an extended period of time to share and debate in-depth technical documents describing the full methodology and risk adjustment model with stakeholders across the private healthcare system.

We welcome the ongoing dialogue and scientific debate on how to measure, report and improve the quality of care provided by private hospitals across SA. Collectively, we must do everything possible to maintain the high standards set in the private healthcare sector and improve the quality of care provided to patients across the system.

Ryan Noach
Chief Executive Officer, Discovery Health
ryann@discovery.co.za

S Afr Med J 2023;113(9):e1139. https://doi.org/10.7196/SAMJ.2023.v113i9.1139

1. Rodseth RN, Smith D, Maslo C, Laubscher A, Thabane L. A critical analysis of Discovery Health's claims-based risk adjustment of mortality rates in South African private sector hospitals. South African Medical Journal 2023;113(1):13-16. https://doi.org/10.7196/SAMJ.2023.v113i1.16768
2. Moodley Naidoo R, Timothy GA, Steenkamp L, Collie S, Greyling MJ. Measuring quality outcomes across hospital systems: Using a claims data model for risk adjustment of mortality rates. S Afr Med J 2019;109(5):299-305. https://doi.org/10.7196/SAMJ.2019.v109i5.13775
3. Lindenauer PK, Bernheim SM, Grady JN, et al. The performance of US hospitals as reflected in risk-standardized 30-day mortality and readmission rates for medicare beneficiaries with pneumonia. J Hosp Med 2010;5(6):e12-18. https://doi.org/10.1002/jhm.8229
4. Tirschwell DL, Longstreth WT Jr. Validating administrative data in stroke research. Stroke 2002;33(10):2465-2470. https://doi.org/10.1161/01.STR.0000032240.28636.BD
5. Bratzler DW, Normand SLT, Wang Y, et al. An administrative claims model for profiling hospital 30-day mortality rates for pneumonia patients. PloS One 2011;6(4):e17401. https://doi.org/10.1371/journal.pone.0017401
6. Krumholz HM, Lin Z, Drye EE, et al. An administrative claims measure suitable for profiling hospital performance based on 30-day all-cause readmission rates among patients with acute myocardial infarction. Circ Cardiovasc Qual Outcomes 2011;4(2):243-252. https://doi.org/10.1161/CIRCOUTCOMES.110.957498
7. Krumholz HM, Wang Y, Mattera JA, et al. An administrative claims model suitable for profiling hospital performance based on 30-day mortality rates among patients with heart failure. Circulation 2006;113(13):1693-1701. https://doi.org/10.1161/circulationaha.105.611194
8. Lindenauer PK, Grosso LM, Wang C, et al. Development, validation, and results of a risk-standardized measure of hospital 30-day mortality for patients with exacerbation of chronic obstructive pulmonary disease. J Hosp Med 2013;8(8):428-435. https://doi.org/10.1002/jhm.2066
9. Butala NM, Strom JB, Faridi KF, et al. Validation of administrative claims to ascertain outcomes in pivotal trials of transcatheter aortic valve replacement. JACC Cardiovasc Interv 2020;13(15):1777-1785. https://doi.org/10.1016/j.jcin.2020.03.049
10. Centers for Medicare & Medicaid Services. 2022 Condition specific mortality measure updates and specificiation reports. CMS, 2022. https://www.cms.gov/files/document/2022-condition-specific-mortality-measures-updates-and-specifications-report.pdf (accessed 15 February 2023).
11. Merative Disease Staging solution brief. Disease staging – analytics for disease identification and corresponding severity. Merative, 2023.
12. Gelman A, Hill J. Data Analysis Using Regression and Multilevel/Hierarchical Models (Analytical Methods for Social Research). Cambridge: Cambridge University Press, 2006. https://doi.org/10.1017/CBO9780511790942
13. Hastie T, Tibshirani R, Friedman JH. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. 2nd ed. New York: Springer, 2009.
14. Shmueli G. To explain or to predict? Statist. Sci 2010;25(3):289-310. https://doi.org/10.1214/10-STS330
15. Breiman L. Random forests. Machine Learning 2001;45:5-32. http://doi.org/10.1023/A:1010933404324