Optum Impact Pro risk algorithm and racial bias (Obermeyer et al.)

IMPLEMENTED 2019 · HEALTH AND LIFE SCIENCES · US · BAND F · PARTLY PROVED
VERSION 1.0 · PUBLISHED 15 NOVEMBER 2026 · UPDATED 15 NOVEMBER 2026 · FACTS CHECKED 15 NOVEMBER 2026

A 2019 Science study found that a widely used commercial risk algorithm, later identified as Optum's Impact Pro, gave Black patients lower risk scores than equally sick White patients because it predicted cost, not illness. New York regulators wrote to UnitedHealth, and Optum worked with the researchers on a revised version.

ABOUT THIS SYNOPSIS

This is MarLuc's summary and grading of publicly available sources, for information only. It is not investment, legal, financial or professional advice, and should not be relied on for any decision. Readers must validate it against the original sources and, where appropriate, challenge its conclusions.

Facts and figures are as reported by the sources listed below, which may be incomplete, self-reported or out of date. Check the original source before relying on any figure.

The impact band, proof marker and SEAM view are MarLuc's opinion. The organisation is not a MarLuc client and has not reviewed or endorsed this synopsis. Names and trademarks belong to their owners.

NOTICE WORDING LAST UPDATED 16 NOVEMBER 2026

AT A GLANCE

What went wrong: AI output or malfunction.

AI used: Predictive machine-learning risk score used to select patients for care-management programmes.

Implemented: 2019. Grade: band F, Failure, Partly proved.

Headline result, as reported: A health risk algorithm that predicted cost under-flagged Black patients for extra care; retraining on cost plus health measures cut the disparity by 84 per cent, per Science News

1 · The case

The facts, as reported by the 5 sources listed at the end. Where a source is silent, the field says "Not reported".

Organisation

Optum (UnitedHealth Group): Impact Pro population health risk algorithm, studied by researchers at UC Berkeley, Brigham and Women's Hospital, Massachusetts General Hospital and the University of Chicago. Commercial risk-prediction algorithm sold by Optum, part of UnitedHealth Group, to US health systems. The researchers analysed its use at one academic

Business challenge

The challenge. US health systems use risk scores to decide which patients receive extra help in high-risk care-management programmes. The algorithm assigned each patient a score based on predicted future cost, which the researchers describe as typical of such tools.

How it was addressed. The researchers used data from one hospital that used Impact Pro (2013 to 2015) and compared patients' risk scores with their chronic illnesses. They shared the finding with the manufacturer, which, per Chicago Booth Review, confirmed it on a national dataset of more than 3.5 million patients and worked with them to test a new label combining cost and health predictions. [1] [2] [3] [4] [5]

Who implemented it

Optum (UnitedHealth Group) (Software vendor): Developed and sells Impact Pro; worked with the researchers on a revised label.
Researchers at UC Berkeley, Brigham and Women's Hospital, Massachusetts General Hospital and the University of Chicago (University or research lab): Audited the algorithm and tested alternative labels.
The hospital that used Impact Pro and supplied the data was not named.

Technology (SEAM pillar E)

SEAM types: T8 Predictive / ML models. Deployment: Not reported. Vendors and products: Optum (UnitedHealth Group), Impact Pro.

Where it was built in: Sold by Optum to health systems; used to score patients for enrolment in high-risk care-management programmes, with a doctor's input for scores from the 55th to 96th percentile.

What changed

Before: the score used predicted cost as a stand-in for health need. After: the researchers and Optum tested a label that combined cost with predicted chronic conditions, which cut the disparity by 84 per cent (Science News). Whether and when a revised Impact Pro was released to customers: Not reported. [1] [2] [3] [4] [5]

Cost of failure

Fines, damages or refunds: Not reported. New York regulators opened an inquiry in October 2019; its outcome: Not reported. Cost of the rework: Not reported.

What it saved

Not applicable [1] [2] [3] [4] [5]

Governance and control

The researchers' audit and the New York letter were outside scrutiny of the tool. Chicago Booth Review says the algorithm was not the only factor in enrolment and that doctors' input applied to patients scoring from the 55th to 96th percentile. Optum said hospitals can address gaps using other factors; the study found the cost label itself carried the bias.

S2 · Accountable owner · Not reported
No accountable owner for the tool's fairness is reported; the researchers did not name the hospital or the company. [2]

A1 · Safe use · Not reported
No policy on safe use of the score is reported.

A2 · Human in the loop · Evidenced
Chicago Booth Review says scores from the 55th to 96th percentile were flagged for possible enrolment depending on the doctor's input, and that the algorithm was not the only factor; the bias still arose in the score itself. [3]

A3 · Monitoring and assurance · Outside scrutiny (a regulator, court, auditor or similar body acted)
Independent researchers audited the tool's scores against chronic illness, and the vendor replicated the analysis on a national dataset. [2] [3]

A4 · Vendor AI · Claimed
UnitedHealth Group said the cost model was highly predictive of cost, which is what it was designed to do, and that health systems can address gaps using other factors. [4]

A5 · Agent controls · Not reported
No agents described.

A6 · Standards and regulation · Outside scrutiny (a regulator, court, auditor or similar body acted)
New York's Department of Financial Services and Department of Health wrote to UnitedHealth Group on 25 October 2019 asking it to demonstrate the algorithm is not discriminatory or stop using it; the inquiry was pending in November 2019. [4] [5]

What worked, what did not, what is unproven

WORKED

– The cost model predicted cost well across races, as the vendor said and the researchers accept: for a given score, future costs were similar for Black and White patients.

– The vendor engaged with the finding, replicated it on a larger dataset and worked with the researchers on a revised label, per press reports.

– A revised label combining cost and chronic-condition predictions cut the disparity by 84 per cent (Science News) or 86 per cent on one measure (Chicago Booth Review).

DID NOT

– At a given risk score, Black patients were considerably sicker than White patients, per the Science abstract.

– Above the 97th percentile, Black patients had 26.3 per cent more chronic illnesses than White patients, per Science News.

– Fewer Black patients were automatically identified for care management: 17.7 per cent against 46.5 per cent without the bias.

UNPROVEN

– How the bias affected patient care: STAT's reporter said in November 2019 that this was not known.

– The outcome of the New York inquiry: Not reported.

– Whether and when customers received a revised version of Impact Pro: Not reported.

Timeline

2013 Start of the 2013 to 2015 patient data later analysed by the researchers [2]

Oct 2019 Obermeyer, Powers, Vogeli and Mullainathan publish their study in Science [1]

Oct 2019 UnitedHealth Group states that the cost model in Impact Pro was highly predictive of cost, which is what it was designed to do [4]

Nov 2019 Optum is reported to be partnering with the researchers; a revised label cuts the disparity by 84 per cent [2]

Nov 2019 STAT reporter says the New York investigation is pending and that Optum is working on a fix [5]

2 · Grade and SEAM assessment

MarLuc's opinion, made on the evidence above. How cases are graded: marluc.co.uk/case-studies/how-we-rate

Level of change and impact

Claimed: Not applicable. Grade: F, Failure. Proof: Partly proved. Level of change: Not graded.

No level assessed: band F, a documented bias in a live commercial risk tool, revised with the vendor. The study is the reference case for proxy-label bias in health algorithms.

Harm is shown in the scores: Black patients needed to be sicker than White patients to receive the same score, so fewer were automatically identified for care-management programmes. It is graded F because the study documents a systematic fault in a live commercial tool that was then revised with the vendor. The marker is partly proved because the score disparity is independently shown and confirmed by the vendor, while the effect on actual care and the outcome of the regulatory inquiry were not reported. The deployment is from before today's generative AI and is a predictive risk model; it is included as the defining example of proxy-label bias.

The SEAM view and lessons learnt

The fault sat in the choice of what the model was asked to predict. Cost looked like a sound proxy for need and was well calibrated across races, yet unequal access to care made it a poor measure of illness. The audit was possible because the hospital held the data used to run the tool. For a mid-sized or larger organisation, the transferable point is to test a model's target variable against the outcome that matters, by group, and to keep access to the data needed to do so.

1. Check what a model predicts against the outcome you care about.
2. Test scores by group before and after rollout.
3. Keep access to vendor model data so that outside audit is possible.

Evidence

Score 3 of 4. Independent: a peer-reviewed study in Science, a regulator's letter, and the vendor's own statement and replication as reported by the press. The full Science text was not available to MarLuc; the paper's abstract was read, and detail is taken from press coverage.

3 · Original sources

Read these in full. To challenge a fact or a grade, use the contact form at marluc.co.uk or the link on https://marluc.co.uk/case-studies/optum-impact-pro-racial-bias-obermeyer

[1] NCBI PubMed record (Science, 2019): Dissecting racial bias in an algorithm used to manage the health of populations
https://eutils.ncbi.nlm.nih.gov/entrez/eutils/efetch.fcgi?db=pubmed&id=31649194&retmode=xml

[2] Science News: A health care algorithm's bias disproportionately hurts black people
https://www.sciencenews.org/article/bias-common-health-care-algorithm-hurts-black-patients

[3] Chicago Booth Review: How Racial Bias Infected a Major Health-Care Algorithm
https://www.chicagobooth.edu/review/how-racial-bias-infected-major-health-care-algorithm

[4] GovTech (Star Tribune, 29 October 2019): NY Regulators Probe for Racial Bias in Health-Care Algorithm
https://www.govtech.com/health/ny-regulators-probe-for-racial-bias-in-health-care-algorithm.html

[5] PBS NewsHour: Racial bias in widely used hospital algorithm, study finds
https://www.pbs.org/newshour/amp/show/racial-bias-in-widely-used-hospital-algorithm-study-finds

Optum (UnitedHealth Group): Impact Pro population health risk algorithm, studied by researchers at UC Berkeley, Brigham and Women's Hospital, Massachusetts General Hospital and the University of Chicago website: https://www.optum.com

Version 1.0 · Published 15 November 2026 · Updated 15 November 2026 · Facts checked against sources 15 November 2026 · Downloaded 26 November 2026. This PDF may be out of date: the latest version, and a way to report an error or a better source, is at https://marluc.co.uk/case-studies/optum-impact-pro-racial-bias-obermeyer

VERSION HISTORY

v1.0 15 November 2026 First published