

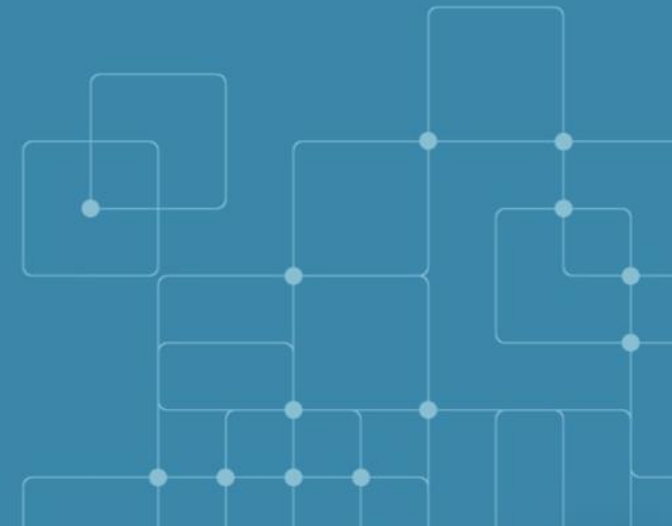
Dr. Soliman Fakeeh Hospital — Jeddah

Value-Based Healthcare & Population Health Management Committee

How healthy do our patients feel?

Case-mix adjustment of the EQ-5D-5L health status score —
comparing our services on equal terms

3,381 patients • 3,491 questionnaires



One question every patient can answer

A knee replacement, a skin condition and a thyroid problem have nothing in common clinically. But every one of those patients can tell us how well they walk, wash, work, hurt and worry. The EQ-5D-5L turns that into a single number, so for the first time we can compare health across the whole hospital on one scale.

1 – It works everywhere

One scale, every department

Disease-specific tools cannot be compared with each other. This one can, which is what makes a hospital-wide view possible at all.

2 – This is what value means

Outcomes that matter to patients

Value-based healthcare defines value as the outcomes that matter to patients for the resources we use. Health status is the broadest of those outcomes.

3 – A number we can act on

A trigger, not a statistic

A patient who reports severe pain or cannot wash themselves has named a problem. The measurement only has value if it changes the next consultation.

The PROM we use: the EQ-5D-5L

Five areas of daily life. For each one the patient picks a level from **1 = no problems** to **5 = unable or extreme**.

What we ask about	Any problem	Severe (4 or 5)
Pain or discomfort	58.3%	8.2%
Anxiety or depression	39.0%	5.1%
Usual activities	34.6%	7.1%
Mobility (walking about)	32.7%	6.9%
Self-care (washing, dressing)	13.0%	3.7%

Pain is by far the commonest problem our patients report, and anxiety or depression is second — ahead of anything to do with movement.

How it becomes one number

The five answers combine into an **index**. **1.00** is full health, **0** is as bad as being dead, and below 0 means a state most people rate as worse than death.

Our hospital average

0.838

Out of 1.00. Patients also rate their health on a 0–100 scale on the day: our average there is **70**.

The clinical picture we place beside it

Every questionnaire was joined, patient by patient, to the clinical and demographic values already held in the hospital system — so the patient's own account of their health and the hospital's record of their biology are read together.

MEASURED VALUES FROM THE RECORD

- **Body mass index**
- **Systolic blood pressure**
- **Diastolic blood pressure**

Recorded by clinicians, not asked of the patient.

PATIENT AND CONDITION FACTORS

- **Age**
- **Sex**
- **Diabetes** recorded
- **Cancer** recorded
- **Autoimmune** condition recorded

ONE THING WE COULD NOT USE

HbA1c was missing for **2,614** of the **3,491** questionnaires — three patients in four.

That is expected: most of these patients do not have diabetes, so no HbA1c was ordered. We left it out of the model rather than drop three quarters of the cohort.

Where the numbers come from

Everything exported for adult DSFH Jeddah patients

3,491 questionnaires · 3,381 patients · 50 departments

All five questions answered and an index value present

3,491 questionnaires — not one was incomplete

Clinically plausible measurements only

3,270 questionnaires · 3,160 patients

Complete clinical record — the cohort we analysed

3,269 questionnaires · 3,159 patients — **94% of everything exported**

What we lost, and why

218 implausible BMI — and **216** of those were a BMI recorded as zero

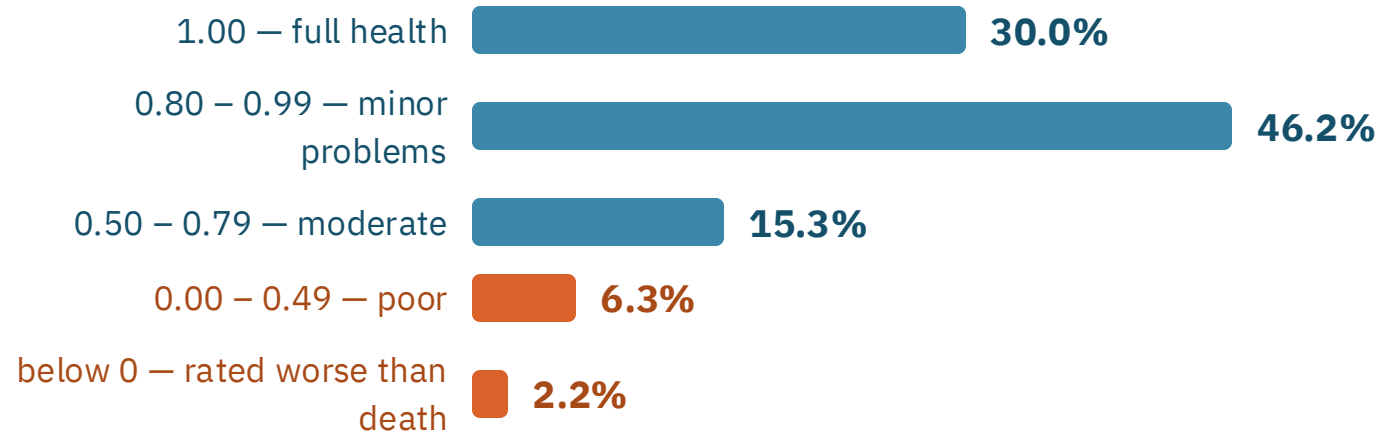
5 implausible or missing blood pressure

A zero BMI is a blank field, not a patient. One data-entry rule would recover all of them.

One patient, one vote

Almost every patient answered once; a few answered up to three times. Each patient still carries a total weight of one, so no one counts twice.

What our patients told us



Bar length is the share of patients in each band of the index. Three quarters of our patients are in good health. The two orange bands — around one patient in twelve — describe lives that are seriously constrained.

15.8%

of our patients report a **severe or extreme problem** in at least one of the five areas — around one in six

And 70% report something

Only **30%** report no problem at all in any area. The average index of **0.838** hides both of those facts, which is why we show the spread.

Why a raw score is not a fair score

A spine clinic and a dermatology clinic do not see the same people. Put their raw averages side by side and you are mostly comparing **who walked through the door**, not the care that was given. With 50 departments on one scale, that problem is larger here than anywhere else we measure.

If we don't adjust

We blame the wrong service

A team doing good work with frail, older, heavier patients looks like a team doing poor work. An improvement plan gets written for a problem that was never theirs.

If we don't adjust

We miss a real problem

A clinic with young, well patients can look excellent while genuinely underperforming. A favourable case mix hides it, and nobody goes looking.

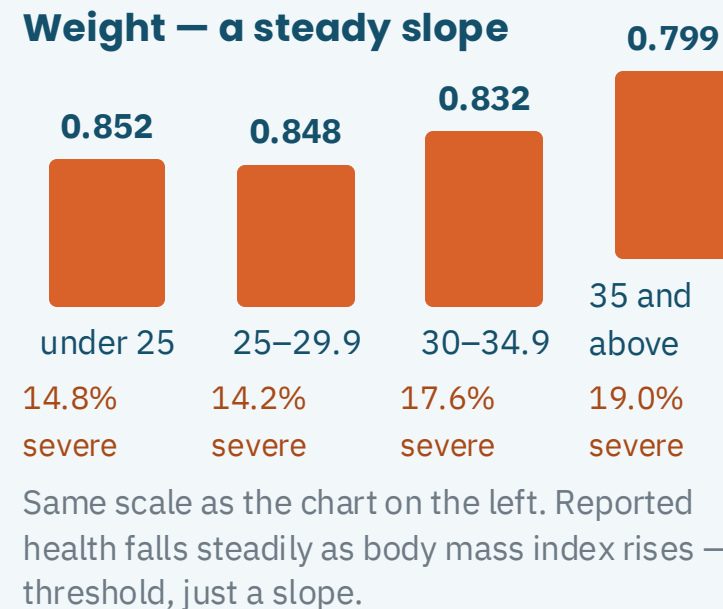
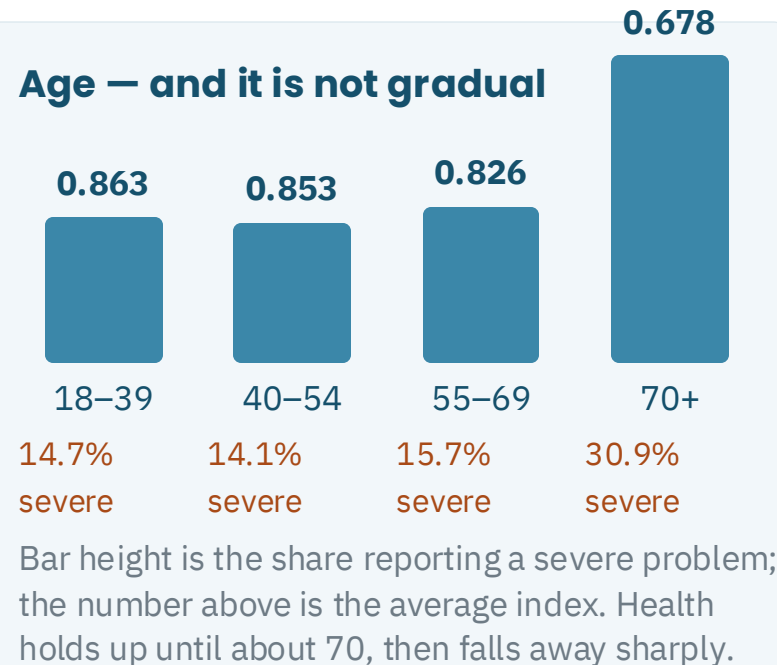
If we don't adjust

We punish taking hard cases

The most damaging outcome. If complex patients lower your score, the quiet incentive is to avoid them. Adjustment removes that incentive.

Case-mix adjustment answers one question: **given the patients this service actually treats, is the result better or worse than we would expect?**

Who reports worse health



What the model confirms

Each 10 years older
-0.033 index points

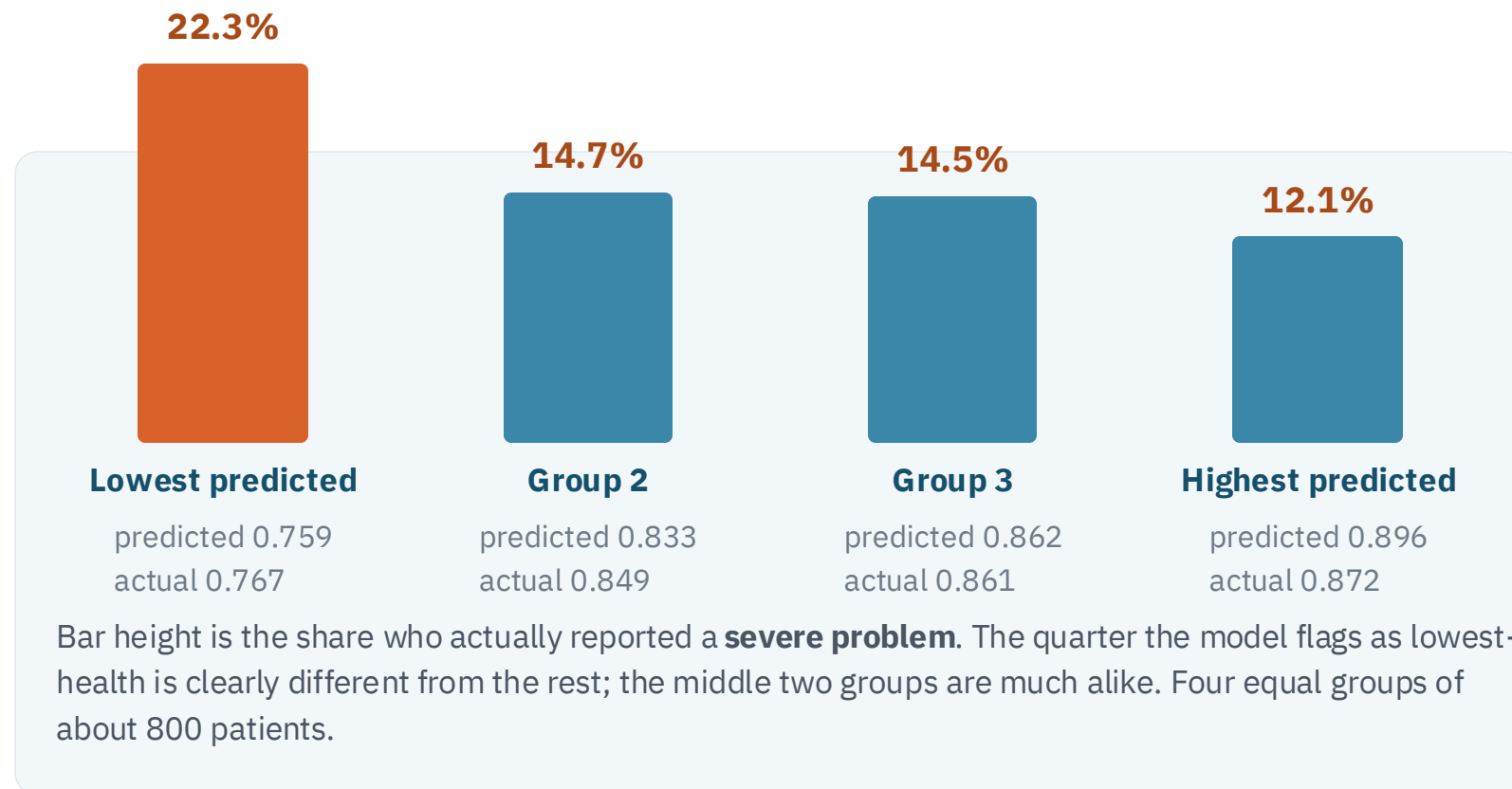
Each 5 BMI units higher
-0.016 index points

Women compared with men
-0.021 index points

Blood pressure and recorded cancer showed no clear effect.
 Bars start at zero.

Risk stratification: does the model find the right patients?

We asked the model to predict a health score for every patient from age, sex, BMI, blood pressure and recorded conditions alone — **without seeing the answers**. Then we sorted patients into four equal groups by that prediction and looked at what they actually reported.



What this means

A patient in the lowest-predicted quarter is **nearly twice** as likely to report a severe problem as one in the highest — 22.3% against 12.1%.

And what it does not mean

It cannot tell you which individual patient is suffering. Even in the lowest group, more than three quarters are not severe. It points the effort; the questionnaire still does the finding.

How the adjustment works

OBSERVED

What this service's patients actually reported

—

EXPECTED

What patients like theirs report across the hospital

+

HOSPITAL AVERAGE

0.838 — the common yardstick for everyone

=

ADJUSTED

The fair, like-for-like number we compare

Worked example — Adult Nephrology

Step	Index
What its 46 patients reported	0.835
What patients like theirs report	– 0.806
Hospital average added back	+ 0.838
Adjusted score	0.866

Its patients are predicted to be in the **worst health of any department**. Judged against that, Nephrology moves from 14th on raw figures to 9th of 22.

The arithmetic always balances

Across the whole hospital, observed, expected and adjusted are identical at 0.838. Adjustment moves health status **between** services; it never invents or removes any.

How sure are we?

Every department's range was built by re-drawing the patient list 1,000 times and refitting the model each time, so the uncertainty shown includes the model's own.

Twenty-two departments, before and after adjustment

Observed Adjusted — like-for-like Hospital average 0.838 · longer bar = better health · scale begins at 0.50



Most departments barely move: the differences between them are real, not an artefact of case mix. The four in orange move by two ranks or more — **Nephrology 14th to 9th** and **Cardiology 12th to 8th** once the difficulty of their patients is allowed for. Each department shown has at least 30 patients. Some names are shortened; the full list of 22 is in the workbook.