A method for calculating the total QALY output and the mean cost / QALY attributable to the health sector

— the case of the UK

Kim Rand, PhD
Maths in Health BV, Klimmen, The Netherlands
Health Services Research Unit, Akershus University Hospital, Lørenskog, Norway
krand@mathsinhealth.com

ISPOR 2026
EE309

Under generous assumptions, the NHS produces QALYs at a mean cost of GBP 33,000–48,500, at or above the upper bound of NICE's new GBP 25,000–35,000 decision range. The true mean cost per attributable QALY is likely higher.

Background

Why this matters

Cost-effectiveness thresholds in HTA are anchored to the marginal cost per QALY of the health sector. The most-cited UK estimate (Claxton 2015) placed this at GBP 12,936. NICE increased their threshold in April 2026 to GBP 20,000 - 30,000. The average cost per QALY produced by the entire health sector a different and more total-budget-relevant quantity has received limited scrutiny.

The displacement gap

Approval of any new technology is cost-effective only if what it displaces has a higher cost per QALY than its own. In practice, displacement tends to happen through aggregate budget envelopes, waiting-list management, and service reconfiguration rather than through systematic identification of the least cost-effective activities. To the extent displacement is poorly targeted, the relevant comparison is the mean cost, not the marginal.

What we add

A transparent, reproducible counterfactual framework for exploring the QALY output attributable to the health sector, and thus the mean cost per QALY, here applied to the UK (NHS 2024, baseline 1948). The method can easily be adjusted to allow for a variety of assumptions to determine plausible ranges of costs per QALY. The methods have been implemented in an interactive web app for exploration. The app and complete code will be released as an R package following full publication.

Methods

A thought experiment: in the absence of the health sector and its longevity and HRQoL benefits, how many QALYs would the current UK population produce? A method of 8 steps.

01 Define attributable QALYs

For each demographic cell d=(age, sex),
$\Delta Q = Q_{obs} - Q_{cf}$. Mean cost / QALY = E / ΔQ.

02 Observed QALYs

$Q_{obs} = \Sigma\, PY \cdot q$. ONS mid-2024 E+W population 61.81 M; ONS Single Year Life Tables 2024; McNamara et al. 2023 norms.

03 The counterfactual challenge

A no-health-sector counterfactual cannot be observed. We anchor in 1948 (NHS founding) mortality and HRQoL. Four sub-problems: different starting population, non-health improvements, HRQoL change, temporal attribution.

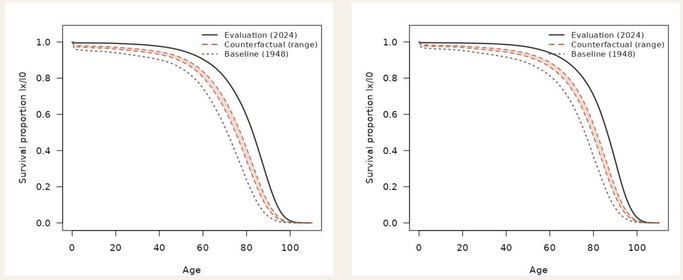

Fig. 1. Survival curves — males (left), females (right). Observed 2024, counterfactual range p_ℓ=0.40–0.60, synthetic baseline 1948.

04 Reconstruct the starting population

$C(x) = N(t_1, x) / [\ell x(t_1) / \ell_0]$. Effective birth-cohort back-calculation sidesteps immigration by treating today's population as cohorts subject to current mortality.

05 pℓ: assumed proportion of longevity gains attributable to the health sector

$\hat{q}x = (1 - p\ell)\cdot qx(t_0) + p\ell \cdot qx(t_1)$. Literature estimates span 0.20–0.73, with most published estimates restricted to amenable-mortality causes. **Generous range reported here is 0.4. – 0.6**, arguably implausibly high.

06 Counterfactual HRQoL via remaining-life utility

Parametric f(k, θ) links EQ-5D utility to remaining years k. Auto-AIC across 5 forms. UK fit: Power (males), Gompertz (females). pℓ-coupled, no separate pH needed. Assumes one year increase in longevity translates to "one year younger" in HRQoL

07 Temporal attribution

Snapshot (lag horizon H = 0) assigns all attributable QALYs to current year. Distributed-lag option uses normalised Koyck weights $wa(k) = \lambda^k / \Sigma_j \lambda^j$ across H forward years. We report H=0 and H=10, λ=0.7.

08 Combine — mean cost per attributable QALY

$\Delta Q = Q_{obs} - Q_{cf}$; mean cost/QALY = E / ΔQ. Ranges reported by treating pℓ and E as range inputs.

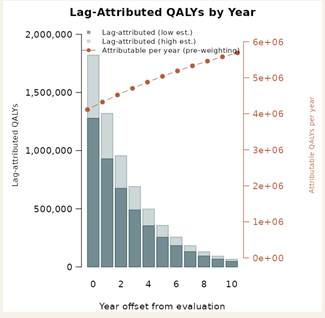

Fig. 2. Distributed-lag attribution. Koyck-weighted bars by year offset; rust overlay = per-year unweighted attributable QALYs.

Results

Setup. Baseline 1948 · evaluation 2024 · ONS mid-2024 E+W 61.81 M · ONS Single Year Life Tables 2024 · McNamara et al. (2023) EQ-5D norms · rate-level counterfactual · endpoint-years (period) method · p_ℓ = 0.40–0.60 · E ∈{£177.4 bn OBR PFD 2023, £181.7 bn HMT PESA 2023–24}.

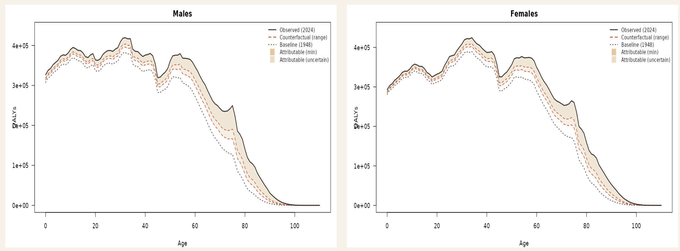

Fig. 3. QALY accrual by age — males (left), females (right). Observed 2024 (charcoal solid), counterfactual range p_ℓ=0.40–0.60 (rust dashed), synthetic 1948 baseline (faint dotted). Gold polygon = attributable surplus.

51.8 M Observed QALYs

3.75 – 5.34 M Attributable QALYs · p_ℓ 0.40–0.60

£33k – 48k Cost / QALY · snapshot

£31k – 44k Cost / QALY · lag H=10

Test your own assumptions live

You don't agree? Have a go!

The app allows user-specification of costs, target years, pℓ , pH, counterfactual methods, time lag horizon, etc. **Note: works best on larger screen.**

https://apps.mathsinhealth.com/apps/meanQALY/

Table 1. Mean cost per QALY (£), snapshot specification.

pℓ	Attributable QALYs	Cost/QALY E=177.4 bn	Cost/QALY E=181.7 bn
0.40	3.75 M	£47,312	£48,458
0.50	4.56 M	£38,871	£39,813
0.60	5.34 M	£33,234	£34,040

Range across pℓ and England-only expenditure: £33,234 – 48,458. Central p_ℓ=0.50, E=£179.55 bn: £39,342 / QALY.

Table 2. Distributed-lag specification (H=10, λ=0.7).

pℓ	Lag-weighted QALYs	E=177.4 bn	E=181.7 bn
0.40	4.10 M	£43,293	£44,342
0.50	4.97 M	£35,693	£36,558
0.60	5.79 M	£30,617	£31,359

Lag-weighted QALYs exceed snapshot by ~9%; cost / QALY falls to £30,617 – 44,342. Qualitatively unchanged.

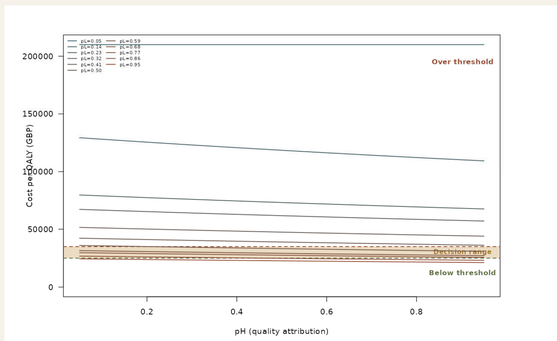

Fig. 4. Cost per QALY vs p_ℓ. E = £177.4 bn, £181.7 bn, and UK-wide hypothetical £271 bn. Horizontal bands mark NICE's old (£20–30k) and new (£25–35k) decision ranges.

Conclusions

1. Fewer QALYs than one might assume

Under headline assumptions the QALY accrual attributable to the health sector is 3.75–5.34 M QALYs, 7.2%–10.3% of total population QALY accrual. The mean cost is roughly £33,000–48,500 per QALY. This sits at or above the upper bound of NICE's new £25,000–35,000 decision range. The health sector therefore delivers substantially less QALY output per pound than is commonly assumed when the marginal estimate is used as a default for displacement.

2. Our assumptions bias results downward

Three reasons: (a) headline pℓ=0.40–0.60 likely over-attributes longevity gains to healthcare; pℓ≈0.25 (amenable-mortality benchmark) is defensible and raises cost/QALY to £62k at pℓ=0.30, £91k at pℓ=0.20; (b) expenditure covers low estimates England only (£177–182 bn) while UK-wide estimates range up to £271 bn, resulting in a downward scope bias on the numerator. Adjusting any of these biases pushes mean cost / QALY upward.

Policy take-away

If displacement from new technology adoption is not highly targeted at the least cost-effective activities, the cost per QALY of displaced services lies between the marginal and the mean. With mean estimates exceeding the marginal, the discussion needs further nuance. High mean costs could reflect inefficiencies, costly care-related activities that are deemed necessary regardless of QALY output, costly infrastructure investments, or that health care is just more expensive than suggested by intuition alone.

CITE AS

Rand K. A method for calculating the total QALY output and the mean cost / QALY attributable to the health sector; the case of the UK. Poster EE309, ISPOR 2026, Philadelphia, PA. *Value in Health* 2026;29(S6).

KEY REFERENCES Claxton K et al. HTA 2015;19(14). · Bunker JP. IJE 2001;30:1260. · Nolte E, McKee CM. BMJ 2003;327:1129. · Naci H et al. Lancet 2025;405:50. · Kind P et al. CHE WP 172, 1999. · Dolan P. Med Care 1997;35:1095. · Barbieri M et al. IJE 2015;44:1549 (HMD). · Sampson C et al. AHEHP 2022;20:651. · Vallejo-Torres L et al. Health Econ 2018;27:746. · Lichtenberg FR. Econ Hum Biol 2022;46:101124.