Decision Integrity for Polygenic Reproductive Reports

A pre-registered comprehension audit

00 The claim

The estimation problem in polygenic prediction is hard, and in part already solved. What no one has measured is comprehension: whether anyone reading a report grasps what a difference of a standard deviation or two from a base rate actually means. Tversky and Kahneman showed decades ago that professionals at a statistics conference fall into the same biased readings of basic numerical inference as everyone else. Fast heuristics mislead us systematically, and experts are not exempt when the framing speaks to intuition rather than calculation. So a report can be statistically correct in every detail and still leave its reader miscalibrated. The worst case is feeling more confident than the variance explained can justify. Choices this consequential deserve, at the minimum, the full use of what the behavioral sciences already know. I have done this before in financial decision-making, where portfolio risk is understood far better through simulation than through a single point estimate. Genomics merits the same: insights aimed squarely at innumeracy, overconfidence, and the misreading of absolute risk differences. I am convinced the field's impact can be lifted, sharply, through genomic decision science.

01 Why this is a gap, not a quibble

Responsible-use language assumes the reader understands the caveats it lists. That assumption is an empirical claim about cognition, and it is almost never tested. Risk communication has thirty years of evidence (Gigerenzer; Reyna's fuzzy-trace theory; the IPDAS decision-aid standards) showing that format governs comprehension at least as much as content does. Polygenic reproductive reports have not been run through that apparatus. The field's ethical defense currently rests on an untested premise about its own readers.

02 Six failure modes the audit targets

01	**Relative vs absolute risk**	Readers import relative risk reduction as absolute. At realistic prevalences and embryo counts the absolute gain is small; a report that leads with relative reduction manufactures demand rather than informs it.
02	**Reference class**	A percentile is undefined without its reference panel, phenotype definition, and time horizon. The denominator is a design choice the reader almost never sees.
03	**Ancestry portability**	Predictive accuracy decays with genetic distance from the discovery cohort. Precision implied for parents outside the discovery ancestry is routinely overstated and rarely quantified.
04	**Within-family attenuation**	Selection among siblings is a within-family contrast; between-family effect sizes inflate the achievable gain. This is the live methodological fault line, and it is invisible at the point of decision.

05	**Rank fragility**	Under score resampling and measurement error, the embryo rank order is unstable. A point-estimate ranking shown without its instability is itself a comprehension failure.
06	**Validity is not utility**	A calibrated score is not a decision. Utility is conditional on absolute risk differences, the alternatives in hand, and what the chooser values. Reports conflate the two.

03 The protocol

Pre-registered, between-subjects, two arms. Readers (prospective parents, and separately clinicians) receive a report and answer items probing each of the six modes. **Arm A** is the report as shipped. **Arm B** is a fact-box redesign holding the underlying numbers fixed. The contrast isolates how much work the *format* is doing, separately from the science.

> Measures: objective comprehension (gist and verbatim) · the gap between subjective and objective understanding (the overconfidence estimate) · decisional conflict · and intended uptake relative to the absolute numbers (the manufactured-demand estimate).

This makes "manufactured demand" a measured construct with a control arm, not an accusation. A report that survives the audit has an empirical defense its competitors cannot claim.

04 Output

A scorecard, the comprehension dataset, and prioritized redesign recommendations. Pre-registration and analysis code public; the underlying report stays proprietary. The deliverable is evidence a partner can put in front of a review board, a regulator, or a skeptic.

05 The ask

Scoping

One design partner, one synthetic or redacted report. Protocol scoped at no charge.

For the study

Coauthorship, data access, study-cost support, or a small contract.

> Non-clinical research, decision-aid design, and comprehension validation. Not genetic counseling. Not medical advice. Not trait-selection advocacy.

— Who is writing this

Stanford PhD, experimental psychology, trained in the Tversky tradition of judgment under uncertainty. Two decades applying behavioral decision science to financial decision-making and product environments, now ported to reproductive genomic risk. The comparative advantage is not statistical genetics; it is the measurement apparatus for comprehension, calibration, and decision-aid validation that the genomics-communication literature has built and that polygenic reports have escaped.