Utility of polygenic embryo screening for disease depends on the selection strategy

Todd Lencz[1,2,3,*], Daniel Backenroth[4], Adam Green[4], Omer Weissbrod[5], Or Zuk[6], and Shai Carmi[4,*]

[1] Departments of Psychiatry and Molecular Medicine, Zucker School of Medicine at Hofstra/Northwell, Hempstead, NY, USA, 11550

[2] Department of Psychiatry, Division of Research, The Zucker Hillside Hospital Division of Northwell Health, Glen Oaks, NY, USA, 11004

[3] Institute for Behavioral Science, The Feinstein Institutes for Medical Research, Manhasset, NY, USA, 11030

[4] Braun School of Public Health and Community Medicine, The Hebrew University of Jerusalem, Jerusalem, Israel, Jerusalem, 9112102

[5] Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA, 02115

[6] Department of Statistics and Data Science, The Hebrew University of Jerusalem, Jerusalem, Israel, 9190501

[*] Corresponding authors: tlencz@northwell.edu (T.L.), shai.carmi@huji.ac.il (S.C.)

Abstract

As of 2019, polygenic risk scores have been utilized to screen *in vitro* fertilization embryos for genetic liability to adult diseases, despite a lack of comprehensive modeling of expected outcomes. In this short report, we demonstrate that a strong determinant of the potential utility of such screening is the *selection strategy* employed, a factor that has not been previously studied. Minimal risk reduction is expected if only extremely high-scoring embryos are excluded, whereas risk reductions are substantially greater if the lowest-scoring embryo (for a given disease) is selected. We systematically examined the relative contributions of the variance explained by the score, the number of embryos, the disease prevalence, and the parental scores on the utility of screening. We discuss the results in the context of relative vs absolute risk, as well as the potential ethical concerns raised by such procedures.

Introduction

Polygenic risk scores (PRS) have become increasingly well-powered, relying on findings from large-scale genome-wide association studies for numerous diseases (Visscher et al., 2017; Wray et al., 2013). The predictive power of a PRS is usually represented by R^2 (or, as we denote it below, r_{ps}^2), the proportion of variance of the liability of the disease explained by the score (Dudbridge, 2013). However, for potential clinical applications, this statistical property needs to be translated into clinically actionable information (Torkamani et al., 2018). This translation requires careful consideration of the specific purposes for which the PRS will be used. For example, individuals with polygenic scores in the top percentiles for coronary artery disease (CAD) were shown to be as likely to have a heart attack as individuals heterozygous for a familial hypercholesterolemia mutation (Khera et al., 2018), and may therefore be good candidates for preventative treatment. By contrast, a wider range of high PRS scores (e.g, top quartile), may provide useful information as a part of a multimodal screening process for cancer risk (Callender et al., 2019; X. Zhang et al., 2018).

Another potential clinical application of PRSs is preimplantation screening of *in vitro* fertilization (IVF) embryos. Although currently in use (Treff, Eccles, et al., 2019) and fraught with ethical concerns (Lázaro-Muñoz et al., 2020), "polygenic embryo screening" (PES) has not yet received much research attention from either geneticists or ethicists. Understanding the statistical properties of PES forms a critical foundation to ethical consideration of the practice (Lázaro-Muñoz et al., 2020). For example, we have recently demonstrated that screening embryos on the basis of polygenic scores for quantitative traits (such as height or intelligence) has limited predictive power in most realistic scenarios (Karavani et al., 2019), and that r_{ps}^2 is a more significant determinant of PES utility for quantitative traits compared with the number of available embryos n. On the other hand, a series of four studies (Lello et al., 2020; Treff, Eccles, et al., 2019; Treff et al., 2020; Treff, Zimmerman, et al., 2019) conducted by a private company providing embryo screening services has suggested that PES for dichotomous disease risk may have significant clinical utility. However, these studies examined a relatively limited range of scenarios, primarily focusing on distinctions between sibling pairs discordant for illness, and did not provide a comprehensive examination of the potential utility of PES.

Here, we use statistical modeling to examine the potential utility of PES for disease risk (defined in terms of relative and absolute risk reduction), comparable to our prior study of PES for quantitative traits (Karavani et al., 2019), with an aim toward informing future ethical deliberations on the practice. We study a range of realistic scenarios, quantifying the role of r_{ps}^2, the number of embryos n, the disease prevalence, and the parental risk scores on the ability of PES to reduce disease risk when screening for a single disease (*Materials and Methods*). We utilize the liability threshold model (Falconer, 1967) to represent disease risk as a continuous liability, comprising genetic and environmental risk factors, under the assumption that individuals with liability exceeding a threshold are affected (*Supplementary Methods*). The liability threshold model was shown to be consistent with data from family-based transmission studies (Wray & Goddard, 2010) and GWAS data (Visscher & Wray, 2015). Consequently, we define the disease risk of a given embryo probabilistically, as the chance that its liability will cross the threshold at any point after birth (**Figure 1A**).

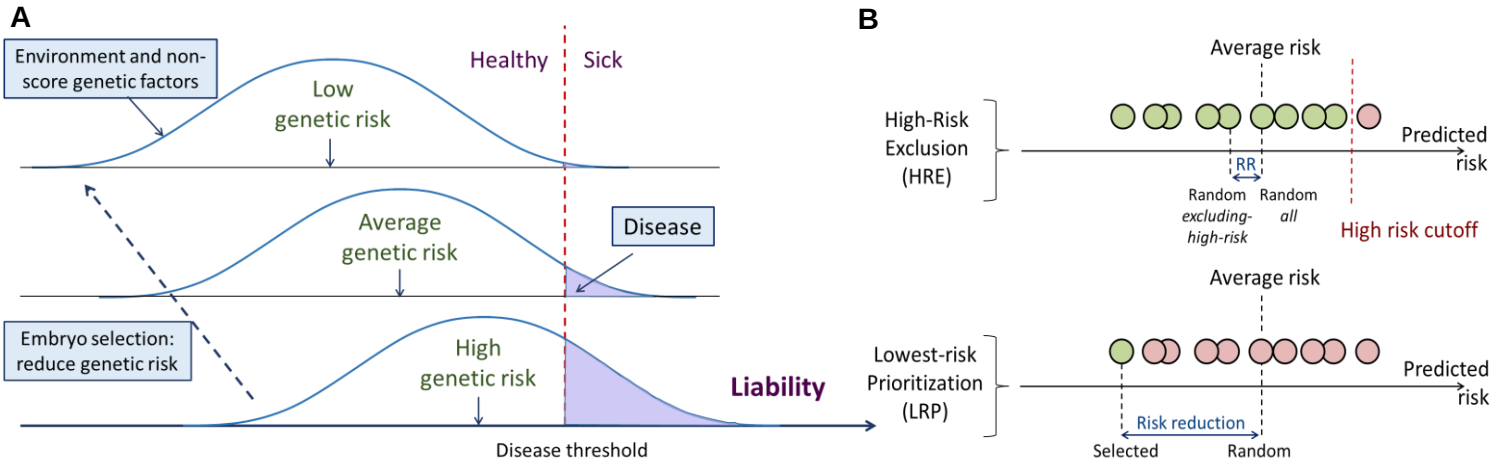


Figure 1. A schematic of the liability threshold model and polygenic embryo screening. (A) An illustration of the liability threshold model (LTM). Under the LTM, it is assumed that each disease has an underlying (unobserved) liability, and that an individual is affected if the total liability is above a threshold. The liability is composed of a genetic component and an environmental component, both assumed to be normally distributed in the population. For a given genetic risk (represented here by the polygenic risk score), the liability is the sum of that risk, plus a normally distributed residual component (environmental + genetic factors not captured by the PRS). Thus, for an individual with high genetic risk (bottom curve), even a modestly elevated (and thus, commonly-occurring) liability-increasing environment will lead to disease. For an individual with low genetic risk (top curve), only an extreme environment will push the liability beyond the disease threshold, making the disease less probable. Thus, disease risk reduction can be achieved with embryo screening by lowering the genetic risk of the implanted embryo. [Note that for the purpose of illustration, panel A displays three discrete levels of genetic risk, although in reality PRS is continuously distributed.] (B) An illustration of the embryo selection strategies considered in this report. In the figure, each embryo is shown as a filled circle, and embryos are sorted based on their predicted risk, i.e., their polygenic risk scores. Excluded embryos are shown in pink, and embryos that can be implanted in green. The risk reduction (RR) is indicated as the difference in risk between a randomly selected embryo (if no polygenic scoring was performed) and the embryo selected based on one of two strategies. In *high-risk exclusion* (HRE), the embryo selected for implantation is random, as long as its PRS is under a high-risk cutoff (usually the top few PRS percentiles). If all embryos are high-risk, a random embryo is selected. In *lowest-risk prioritization* (LRP), the embryo with the lowest PRS is selected for implantation. As we describe below, the LRP strategy yields much larger disease risk reductions.

In studies of potential clinical utility of PRSs in adults, attention has focused on those in the highest percentiles of risk, in which odds ratios become sufficiently large to be clinically meaningful (Chatterjee et al., 2016; Dai et al., 2019; Gibson, 2019; Khera et al., 2018; Mars et al., 2020; Mavaddat et al., 2019; Torkamani et al., 2018). To our knowledge, PES is currently being employed in a similar fashion, following a strategy of excluding embryos with extremely high (top 2-percentiles) PRS (Treff, Eccles, et al., 2019; Treff et al., 2020), which we term "*high-risk exclusion*" (HRE: **Figure 1B**, upper panel). In such a strategy, after high-risk embryos are set aside, an embryo is randomly selected for implantation among the remaining available embryos. (In the rare scenario that all embryos are high-risk, we assume a random embryo is selected among them.) We utilized both theory and simulations to study this strategy, given varying disease prevalence, strength of PRS, and embryo exclusion thresholds. As

elaborated below, we demonstrate that the HRE strategy has very limited utility, but that a strategy of selecting the embryo with the lowest genetic risk can result in large risk reductions.

Results and Discussion

We quantified the outcome of PES in terms of relative risk reduction (RRR), defined as $\mathrm{RRR} = \frac{K - P(\mathrm{disease})}{K} = 1 - \frac{P(\mathrm{disease})}{K}$, where K is the prevalence. For example, if a disease has prevalence of 5% and the selected embryo has a probability of 3% to be affected, the relative risk reduction is 40%. The achievable risk reductions with the *high-risk exclusion* strategy (*Materials and Methods*) are plotted in **Figure 2** (upper row), showing strong dependence on the PRS exclusion threshold. When the 2-percentile threshold is applied (straight black lines), the reduction in risk is limited; RRR is <10% in all scenarios where $r_{ps}^2 \leq 0.1$. Currently, $r_{ps}^2 \approx 0.1$ is the upper limit of predictive power (on the liability scale) of PRS for most complex diseases (Khera et al., 2018; Lambert et al., 2019), with the exception of a few disorders with large-effect common variants (such as Alzheimer's disease or type 1 diabetes) (Sharp et al., 2019; Q. Zhang et al., 2020).

In the future, more accurate PRSs are expected; however, it has been suggested that $r_{ps}^2 = 0.3$, which is at the top end of the common-variant SNP heritability for even the most heritable diseases such as schizophrenia and celiac disease (Holland et al., 2020; Y. Zhang et al., 2018), is the maximal realistic value for the foreseeable future (Wray et al., 2020). At this value, relative risk reduction would only be 20% for $K = 0.01$, 9% for $K = 0.05$, and 3% for $K = 0.2$. These small gains achieved with *high-risk exclusion* follow because the overwhelming majority of affected individuals do not have extreme scores (Murray et al., 2020; Wald & Old, 2019).

Risk reduction increases as the threshold for exclusion is expanded to include the top quartile of scores, and then reaches a maximum at ≈25-50% under a range of prevalence and r_{ps}^2 values. For all of these simulations, we set the number of available embryos to $n = 5$, which is a reasonable estimate of the number of testable, viable embryos from a typical IVF cycle (Sunkara et al., 2011). Simulations show that these estimates do not change much with increasing the number of embryos (see **Supplementary Figure 1**). This holds especially at more extreme threshold values, since most batches of n embryos will not contain any embryos with a PRS within, e.g., the top 2-percentiles.

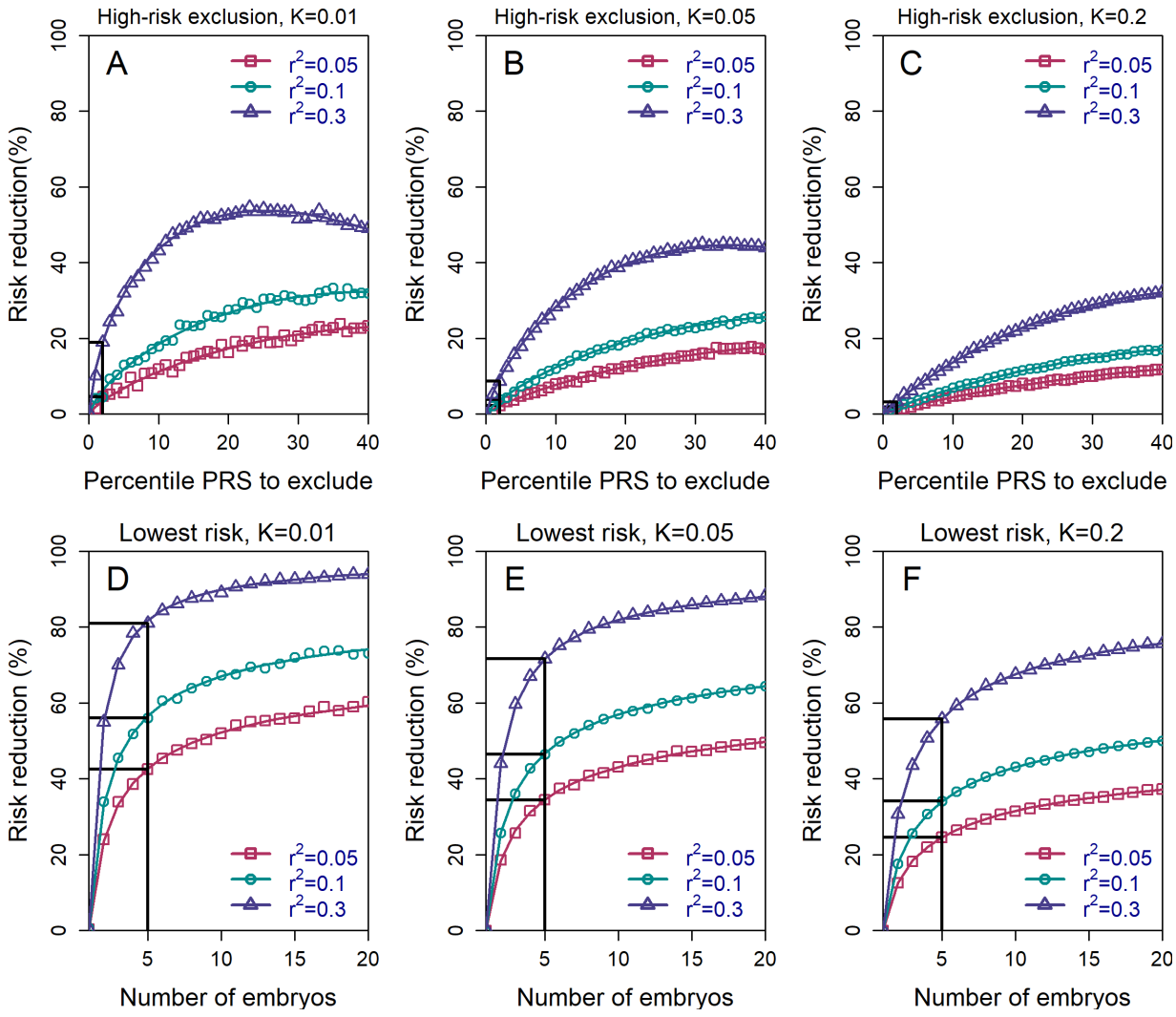
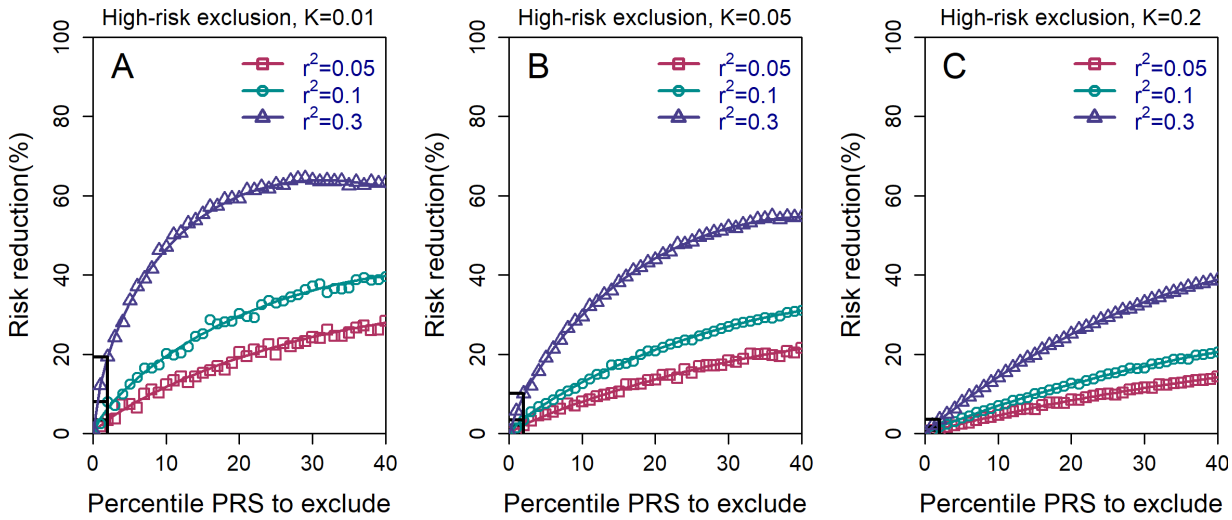


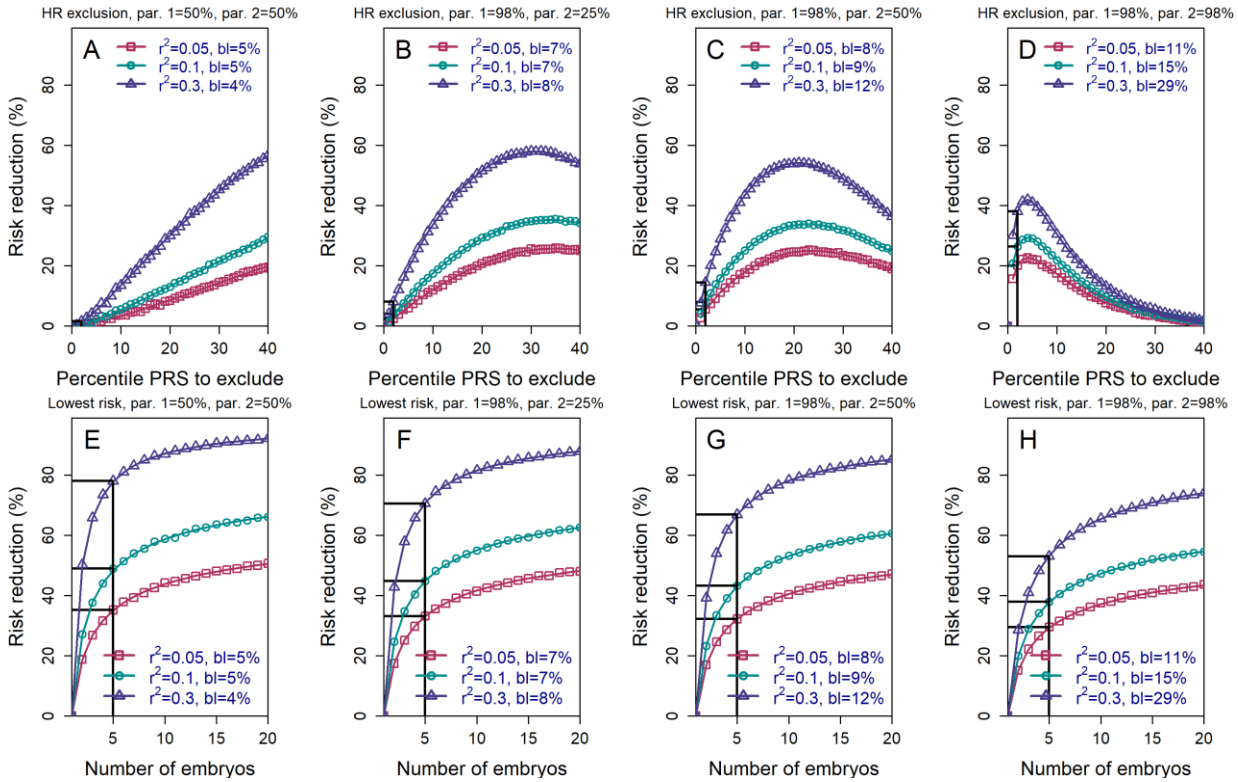
Figure 2. The relative risk reduction across selection strategies and disease parameters. The relative risk reduction (RRR) is defined as $\big(K - P(\text{disease})\big)/K$, where K is the disease prevalence, and $P(\text{disease})$ is the probability of the implanted embryo to become affected. The RRR is shown for the *high-risk exclusion* (HRE) strategy in the upper row (panels A-C), and for the *lowest-risk prioritization* (LRP) in the lower row (panels D-F). See **Figure 1** for the definitions of the strategies. Results are shown for values of $K = 1\%, 5\%, 20\%$ in each panel, and within each panel, for variance explained by the PRS (on the liability scale) $r_{ps}^2 = 5\%, 10\%, 30\%$ (legends). Symbols denote the results of Monte-Carlo simulations (*Materials and Methods*), where PRSs of embryos were drawn based on a multivariate normal distribution, assuming PRSs are standardized to have zero mean and variance r_{ps}^2, and accounting for the genetic similarity between siblings (Eq. (4) in the *Supplementary Methods*). In each simulated set of n sibling embryos ($n = 5$ for all simulations under HRE), one embryo was selected according to the selection strategy. The liability of the selected embryo was computed by adding a residual component (drawn from a normal distribution with zero mean and variance $1 - r_{ps}^2$) to its standardized polygenic score. The embryo was considered affected if its liability exceeded z_K, the (upper) K-quantile of the standard normal distribution. We repeated the simulations over 10^6 sets of embryos and computed the disease risk. In each panel, curves correspond to theory: Eq. (25) in the *Supplementary Methods* for the HRE strategy, and Eq. (14) of the *Supplementary Methods* for the LRP strategy. Black straight lines correspond to the RRR achieved when excluding embryos at the top 2% of the PRS (for HRE, upper panels) or for selecting the lowest risk embryo out of $n = 5$ (for LRP, lower panels).



Supplementary Figure 1. The relative risk reduction for the *high-risk exclusion* strategy, with $n = 10$ available embryos. All details are exactly as in panels (A-C) in **Figure 2** of the main text, except that we simulated $n = 10$ embryos.

The *high-risk exclusion* strategy treats non-high-risk embryos equally; however, other strategies are possible. For example, current research on optimizing IVF protocols focuses on ranking embryos for potential viability, on the basis of microscopy and time-lapse imaging (Bormann et al., 2020; Montag et al., 2013; Rhenman et al., 2015). Thus, it is readily conceivable that the ranking of embryos on the basis of disease PRS could also be attempted. We term the implantation of the embryo with the lowest PRS as "*lowest-risk-prioritization*" (LRP; **Figure 1b**, lower panel). As we show in **Figure 2** (lower panels), for LRP with $n = 5$ available embryos, RRR>20% across the entire range of prevalence and r_{ps}^2 parameters considered, and can reach ≈50% for $K \leq 5\%$ and $r_{ps}^2 = 0.1$, and even ≈80% for $K = 1\%$ and $r_{ps}^2 = 0.3$. While RRR continues to increase as the number of available embryos increases, the gains are quickly diminishing after $n = 5$.

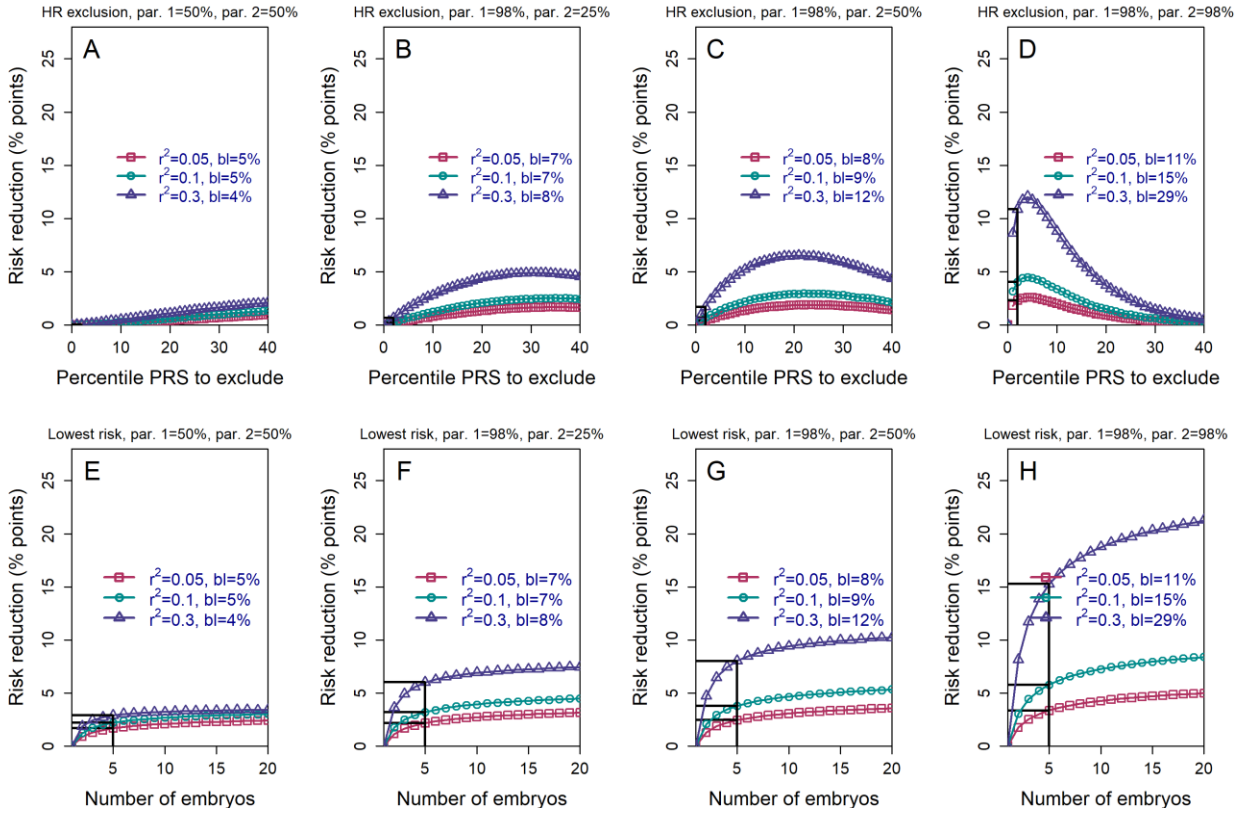
We also examined the effects of parental PRSs on the achievable risk reduction, given the possibility that families with high genetic risk for a given disease would be more likely to seek PES. **Supplementary Figure 2** demonstrates that, as expected, the HRE strategy shows greater relative risk reduction as parental PRS increases, in particular when excluding only very high-scoring embryos. In contrast, the RRR for the LRP strategy is relatively stable across parental PRSs. Nevertheless, the RRR for the LRP strategy remains greater than that for the HRE strategy across all possible combinations of parameters.



Supplementary Figure 2. The relative risk reduction when the polygenic risk scores of the parents are known. Panels (A)-(D) are for the *high-risk exclusion* (HRE) strategy, while panels (E)-(H) are for the *lowest-risk prioritization* (LRP) strategy. All details are as in **Figure 2** of the main text, except the following. First, we fixed the prevalence at $K = 5\%$. Second, in the simulations, we drew the PRS of each embryo as $s_i = x_i + c$ $(i = 1, \dots, n)$, where x_i is an embryo-specific component (independent across embryos) and c is the shared component, also representing the mean parental PRS (*Materials and Methods*). This is so far as in **Figure 2**; however, here we assumed that c is given, equal to the average PRSs of the two parents. In each panel, we consider a different pair of PRSs for the parents. For example, in panels (A) and (E), both parents ("par. 1" and "par. 2") have PRS equal to the 50% percentile of the PRS distribution; in panels (B) and (F), one parent has PRS equal to the 98% percentile of the PRS distribution, while the other has PRS equal to the 25% percentile; and so on. Third, in the simulations, we computed the risk reduction (according to either strategy) relative to a baseline, obtained from the same sets of simulations, when we always selected the first embryo. The baseline risk is indicated in each legend as "bl". Note that the baseline risk depends on the variance explained by the PRS, because the parental PRSs are determined as percentiles of the population distribution of the score, which has variance r_{ps}^2. Finally, we computed the theoretical disease risk for the HRE strategy using Eq. (23) of the *Supplementary Methods*, the disease risk for the LRP strategy using Eq. (17), and the relative risk reduction (shown in curves) for both strategies using Eq. (30).

It is important to emphasize that all of the above results are presented in terms of *relative* risk reduction. The clinical interpretation of these changes in terms of *absolute risk* will vary based on the population prevalence of the disorder, and can offer a very different perspective on the magnitude of the effects (**Supplementary Figure 3)** (Gordis, 2014; Lázaro-Muñoz et al., 2020; Murray et al., 2020). Specifically, a large *relative* risk reduction may result in a very small change in *absolute* risk for a rare disease. As an example, schizophrenia is a highly heritable (Sullivan et al., 2003), serious mental illness with prevalence of at most 1% (Perälä et al., 2007). The most recent large-scale GWAS meta-analysis for

schizophrenia (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2020) has reported that a PRS accounts for approximately 8% of the variance on the liability scale. Our model shows that a 52% RRR is attainable using the LRP strategy with $n = 5$ embryos. However, this translates to only a ≈0.5 percentage point reduction on the absolute scale: a randomly-selected embryo would have a 99% chance of not developing schizophrenia, compared to a 99.5% chance for an embryo selected according to LRP. In the case of a more common disease such as type 2 diabetes, with a lifetime prevalence in excess of 20% in the United States (Geiss et al., 2014), the RRR with $n = 5$ embryos (if the full SNP heritability of 17% (Y. Zhang et al., 2018) were achieved) is 43%, which would correspond to a >8 percentage point reduction in absolute risk.



Supplementary Figure 3. The *absolute* risk reduction when the polygenic risk scores of the parents are known. All details are the same as in **Supplementary Figure 2**, except that the *absolute* (rather than the relative) risk reduction is shown. The absolute risk reduction is defined as the difference between the baseline disease risk (given the parental PRSs; legends) and the risk following either strategy of embryo selection. It is plotted as percentage points.

The results of the present study demonstrate that, contrary to our previous study of PES for quantitative traits (Karavani et al., 2019), substantial effects of PES for disease risk are attainable under certain conditions. Specifically, we observed that the *selection strategy* is a crucial determinant of risk reduction. The use of PRS in adults has focused on those at highest risk (Chatterjee et al., 2016; Dai et al., 2019; Gibson, 2019; Khera et al., 2018; Mars et al., 2020; Mavaddat et al., 2019; Torkamani et al., 2018), for whom there may be maximal clinical benefit of screening and intervention. However, as PRSs

have relatively low sensitivity, such a strategy is relatively ineffective in reducing the overall population disease burden (Ala-Korpela & Holmes, 2020; Wald & Old, 2019). Similarly, in the context of PES, exclusion of high-risk embryos will result in relatively modest risk reductions.

The differential performance of PES across selection strategies may exacerbate the already profound ethical issues raised by PES (as we have recently reviewed (Lázaro-Muñoz et al., 2020)). These issues include stigmatization (McCabe & McCabe, 2011), autonomy (including "choice overload" (Hadar & Sood, 2014)), and equity (Sueoka, 2016). In addition, the ever-present specter of eugenics (Lombardo, 2018) may be especially salient in the context of the LRP strategy. These thorny issues, as well as the difference between relative and absolute risk reduction, may also be difficult to communicate to couples seeking assisted reproductive technologies (Cunningham et al., 2015; Wilkinson et al., 2019). We thus call for urgent deliberations amongst key stakeholders (including researchers, clinicians, and patients) to address governance of PES and for the development of policy statements by professional societies. We hope that our statistical framework can provide an empirical foundation for these critical ethical and policy deliberations.

The present study has several limitations. First, our analytical modeling is based on several simplifying assumptions, though, as we discuss in the *Supplementary Methods*, we do not expect them to qualitatively impact the risk reduction estimates. Second, in practice, it is possible to profile an embryo on the basis of multiple disease PRSs simultaneously, which may affect the overall and per-disease outcomes. Particularly, given the fact that polygenic risk scores are often correlated across diseases (Watanabe et al., 2019; Zheng et al., 2017), selecting based on the PRS of one disease may increase or decrease risk for other diseases. Third, we did not model the outcomes of embryo screening conditional on the disease status of parents or other relatives. These more complex questions will be the subject of future work. Finally, the accuracy of PRSs may be sub-optimal when applied in non-European populations, across different socio-economic groups, or within siblings (as compared to between unrelated individuals) (Duncan et al., 2019; Mostafavi et al., 2020). Moreover, currently constructed PRS do not include rare or ultra-rare variants (including copy number variants), which may have greater penetrance than common variants. For the purposes of our calculations, we assume r_{ps}^2 represents the realistic accuracy achievable (within-family) in a real-world setting in the target population.

Materials and Methods

Our model for the polygenic risk scores of a batch of n IVF embryos is described in detail in the *Supplementary Methods*. Briefly, we write the polygenic scores of the embryos as $(s_1, \dots, s_n)$ where $s_i = x_i + c$, $(x_1, \dots, x_n)$ are embryo-specific independent variables with distribution $x_i \sim N\left(0, r_{ps}^2/2\right)$ (where r_{ps}^2 is the proportion of variance in liability explained by the score), and c is a shared component with distribution $c \sim N\left(0, r_{ps}^2/2\right)$, also representing the mean parental score. In each batch, an embryo is selected according to the selection strategy. For *lowest-risk prioritization*, we select the embryo with the lowest value of s. For *high-risk exclusion*, we select the first embryo with score $s < z_q r_{ps}$, where z_q is the $(1-q)$-quantile of the standard normal distribution. (If no such embryo exists, we select the first embryo.) The liability of the selected embryo is computed as $y = s + e$, where $e \sim N\left(0, 1 - r_{ps}^2\right)$. We

designate the embryo as affected if $y > z_K$, where z_K is the $(1-K)$-quantile of the standard normal distribution and K is the prevalence. For each parameter setting, we compute the disease probability as the fraction of batches (out of 10^6 repeats) in which the selected embryo was affected.

We also compute the disease probability analytically. For *lowest-risk prioritization*, we derive the disease risk based on the theory of order statistics. For *high-risk exclusion*, we first condition on the shared component c, and then study separately the case when all embryos are high-risk (have score $s > z_q r_{ps}$), in which the distribution of the unique component of the selected embryo (x) is a normal variable truncated from below at $z_q r_{ps}$, and the case when at least one embryo has score $s < z_q r_{ps}$, in which x is a normal variable truncated from above. We solved the integrals in the final expressions numerically in R. We also provide explicit results for the case when the family-specific component c is known. For full details and for an additional discussion of previous work and limitations, see the *Supplementary Methods*.

Acknowledgements

We thank Gabriel Lázaro-Muñoz and Stacey Pereira for discussions.

References

Ala-Korpela, M., & Holmes, M. V. (2020). Polygenic risk scores and the prediction of common diseases. *International Journal of Epidemiology*, *49*(1), 1–3. https://doi.org/10.1093/ije/dyz254

Bormann, C. L., Kanakasabapathy, M. K., Thirumalaraju, P., Gupta, R., Pooniwala, R., Kandula, H., Hariton, E., Souter, I., Dimitriadis, I., Ramirez, L. B., Curchoe, C. L., Swain, J., Boehnlein, L. M., & Shafiee, H. (2020). Performance of a deep learning based neural network in the selection of human blastocysts for implantation. *ELife*, *9*, e55301. https://doi.org/10.7554/eLife.55301

Callender, T., Emberton, M., Morris, S., Eeles, R., Kote-Jarai, Z., Pharoah, P. D. P., & Pashayan, N. (2019). Polygenic risk-tailored screening for prostate cancer: A benefit-harm and cost-effectiveness modelling study. *PLoS Medicine*, *16*(12), e1002998. https://doi.org/10.1371/journal.pmed.1002998

Chatterjee, N., Shi, J., & García-Closas, M. (2016). Developing and evaluating polygenic risk prediction models for stratified disease prevention. *Nature Reviews. Genetics*, *17*(7), 392–406. https://doi.org/10.1038/nrg.2016.27

Cunningham, J., Goldsmith, L., & Skirton, H. (2015). The evidence base regarding the experiences of and attitudes to preimplantation genetic diagnosis in prospective parents. *Midwifery*, *31*(2), 288–296. https://doi.org/10.1016/j.midw.2014.09.010

Dai, J., Lv, J., Zhu, M., Wang, Y., Qin, N., Ma, H., He, Y.-Q., Zhang, R., Tan, W., Fan, J., Wang, T., Zheng, H., Sun, Q., Wang, L., Huang, M., Ge, Z., Yu, C., Guo, Y., Wang, T.-M., … Shen, H. (2019). Risk loci identification and polygenic risk score in prediction of lung cancer: a large-scale prospective cohort study in Chinese. *The Lancet. Respiratory Medicine*, *7*(10), 881–891. https://doi.org/10.1016/S2213-2600(19)30144-4

Dudbridge, F. (2013). Power and predictive accuracy of polygenic risk scores. *PLoS Genetics*, *9*(3), e1003348. https://doi.org/10.1371/journal.pgen.1003348

Duncan, L., Shen, H., Gelaye, B., Meijsen, J., Ressler, K., Feldman, M., Peterson, R., & Domingue, B. (2019). Analysis of polygenic risk score usage and performance in diverse human populations. *Nature Communications*, *10*(1), 3328. https://doi.org/10.1038/s41467-019-11112-0

Falconer, D. S. (1967). The inheritance of liability to diseases with variable age of onset, with particular reference to diabetes mellitus. *Annals of Human Genetics*, *31*(1), 1–20. https://doi.org/10.1111/j.1469-1809.1967.tb01249.x

Geiss, L. S., Wang, J., Cheng, Y. J., Thompson, T. J., Barker, L., Li, Y., Albright, A. L., & Gregg, E. W. (2014). Prevalence and incidence trends for diagnosed diabetes among adults aged 20 to 79 years, United States, 1980-2012. *JAMA*, *312*(12), 1218–1226. https://doi.org/10.1001/jama.2014.11494

Gibson, G. (2019). On the utilization of polygenic risk scores for therapeutic targeting. *PLoS Genetics*, *15*(4), e1008060. https://doi.org/10.1371/journal.pgen.1008060

Gordis, L. (2014). *Epidemiology*.

Hadar, L., & Sood, S. (2014). When knowledge is demotivating: subjective knowledge and choice overload. *Psychological Science*, *25*(9), 1739–1747. https://doi.org/10.1177/0956797614539165

Holland, D., Frei, O., Desikan, R., Fan, C.-C., Shadrin, A. A., Smeland, O. B., Sundar, V. S., Thompson, P., Andreassen, O. A., & Dale, A. M. (2020). Beyond SNP heritability: Polygenicity and discoverability of phenotypes estimated with a univariate Gaussian mixture model. *PLoS Genetics*, *16*(5), e1008612. https://doi.org/10.1371/journal.pgen.1008612

Karavani, E., Zuk, O., Zeevi, D., Barzilai, N., Stefanis, N. C., Hatzimanolis, A., Smyrnis, N., Avramopoulos, D., Kruglyak, L., Atzmon, G., Lam, M., Lencz, T., & Carmi, S. (2019). Screening Human Embryos for Polygenic Traits Has Limited Utility. *Cell*, *179*(6), 1424-1435.e8. https://doi.org/10.1016/j.cell.2019.10.033

Khera, A. V., Chaffin, M., Aragam, K. G., Haas, M. E., Roselli, C., Choi, S. H., Natarajan, P., Lander, E. S., Lubitz, S. A., Ellinor, P. T., & Kathiresan, S. (2018). Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nature Genetics*, *50*(9), 1219–1224. https://doi.org/10.1038/s41588-018-0183-z

Lambert, S. A., Abraham, G., & Inouye, M. (2019). Towards clinical utility of polygenic risk scores. *Human Molecular Genetics*, *28*(R2), R133–R142. https://doi.org/10.1093/hmg/ddz187

Lázaro-Muñoz, G., Pereira, S., Carmi, S., & Lencz, T. (2020). Screening embryos for polygenic conditions and traits: ethical considerations for an emerging technology. *Genetics in Medicine: Official Journal of the American College of Medical Genetics*. https://doi.org/10.1038/s41436-020-01019-3

Lello, L., Raben, T. G., & Hsu, S. D. H. (2020). Sibling validation of polygenic risk scores and complex trait prediction. *Scientific Reports*, *10*(1), 13190. https://doi.org/10.1038/s41598-020-69927-7

Lombardo, P. A. (2018). The power of heredity and the relevance of eugenic history. *Genetics in Medicine*, *20*(11), 1305–1311. https://doi.org/10.1038/s41436-018-0123-4

Mars, N., Koskela, J. T., Ripatti, P., Kiiskinen, T. T. J., Havulinna, A. S., Lindbohm, J. V., Ahola-Olli, A., Kurki, M., Karjalainen, J., Palta, P., FinnGen, Neale, B. M., Daly, M., Salomaa, V., Palotie, A., Widén, E., & Ripatti, S. (2020). Polygenic and clinical risk scores and their impact on age at onset and prediction of cardiometabolic diseases and common cancers. *Nature Medicine*, *26*(4), 549–557. https://doi.org/10.1038/s41591-020-0800-0

Mavaddat, N., Michailidou, K., Dennis, J., Lush, M., Fachal, L., Lee, A., Tyrer, J. P., Chen, T.-H., Wang, Q., Bolla, M. K., Yang, X., Adank, M. A., Ahearn, T., Aittomäki, K., Allen, J., Andrulis, I. L., Anton-Culver, H., Antonenkova, N. N., Arndt, V., … Easton, D. F. (2019). Polygenic Risk Scores for Prediction of Breast Cancer and Breast Cancer Subtypes. *American Journal of Human Genetics*, *104*(1), 21–34. https://doi.org/10.1016/j.ajhg.2018.11.002

McCabe, L. L., & McCabe, E. R. B. (2011). Down syndrome: coercion and eugenics. *Genetics in Medicine: Official Journal of the American College of Medical Genetics*, *13*(8), 708–710. https://doi.org/10.1097/GIM.0b013e318216db64

Montag, M., Toth, B., & Strowitzki, T. (2013). New approaches to embryo selection. *Reproductive Biomedicine Online*, *27*(5), 539–546. https://doi.org/10.1016/j.rbmo.2013.05.013

Mostafavi, H., Harpak, A., Agarwal, I., Conley, D., Pritchard, J. K., & Przeworski, M. (2020). Variable prediction accuracy of polygenic scores within an ancestry group. *ELife*, *9*. https://doi.org/10.7554/eLife.48376

Murray, G. K., Lin, T., Austin, J., McGrath, J. J., Hickie, I. B., & Wray, N. R. (2020). Could Polygenic Risk Scores Be Useful in Psychiatry?: A Review. *JAMA Psychiatry*. https://doi.org/10.1001/jamapsychiatry.2020.3042

Perälä, J., Suvisaari, J., Saarni, S. I., Kuoppasalmi, K., Isometsä, E., Pirkola, S., Partonen, T., Tuulio-Henriksson, A., Hintikka, J., Kieseppä, T., Härkänen, T., Koskinen, S., & Lönnqvist, J. (2007). Lifetime prevalence of psychotic and bipolar I disorders in a general population. *Archives of General Psychiatry*, *64*(1), 19–28. https://doi.org/10.1001/archpsyc.64.1.19

Rhenman, A., Berglund, L., Brodin, T., Olovsson, M., Milton, K., Hadziosmanovic, N., & Holte, J. (2015). Which set of embryo variables is most predictive for live birth? A prospective study in 6252 single embryo transfers to construct an embryo score for the ranking and selection of embryos. *Human Reproduction (Oxford, England)*, *30*(1), 28–36. https://doi.org/10.1093/humrep/deu295

Schizophrenia Working Group of the Psychiatric Genomics, C., Ripke, S., Walters, J. T., & O'Donovan, M. C. (2020). Mapping genomic loci prioritises genes and implicates synaptic biology in schizophrenia. *MedRxiv*, 2020.09.12.20192922. https://doi.org/10.1101/2020.09.12.20192922

Sharp, S. A., Rich, S. S., Wood, A. R., Jones, S. E., Beaumont, R. N., Harrison, J. W., Schneider, D. A., Locke, J. M., Tyrrell, J., Weedon, M. N., Hagopian, W. A., & Oram, R. A. (2019). Development and Standardization of an Improved Type 1 Diabetes Genetic Risk Score for Use in Newborn Screening and Incident Diagnosis. *Diabetes Care*, *42*(2), 200–207. https://doi.org/10.2337/dc18-1785

Sueoka, K. (2016). Preimplantation genetic diagnosis: an update on current technologies and ethical considerations. *Reproductive Medicine and Biology*, *15*(2), 69–75. https://doi.org/10.1007/s12522-015-0224-6

Sullivan, P. F., Kendler, K. S., & Neale, M. C. (2003). Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies. *Archives of General Psychiatry*, *60*(12), 1187–1192. https://doi.org/10.1001/archpsyc.60.12.1187

Sunkara, S. K., Rittenberg, V., Raine-Fenning, N., Bhattacharya, S., Zamora, J., & Coomarasamy, A. (2011). Association between the number of eggs and live birth in IVF treatment: an analysis of 400 135 treatment cycles. *Human Reproduction (Oxford, England)*, *26*(7), 1768–1774. https://doi.org/10.1093/humrep/der106

Torkamani, A., Wineinger, N. E., & Topol, E. J. (2018). The personal and clinical utility of polygenic risk scores. *Nature Reviews. Genetics*, *19*(9), 581–590. https://doi.org/10.1038/s41576-018-0018-x

Treff, N. R., Eccles, J., Lello, L., Bechor, E., Hsu, J., Plunkett, K., Zimmerman, R., Rana, B., Samoilenko, A., Hsu, S., & Tellier, L. C. A. M. (2019). Utility and First Clinical Application of Screening Embryos for Polygenic Disease Risk Reduction. *Frontiers in Endocrinology*, *10*, 845. https://doi.org/10.3389/fendo.2019.00845

Treff, N. R., Eccles, J., Marin, D., Messick, E., Lello, L., Gerber, J., Xu, J., & Tellier, L. C. A. M. (2020). Preimplantation Genetic Testing for Polygenic Disease Relative Risk Reduction: Evaluation of Genomic Index Performance in 11,883 Adult Sibling Pairs. *Genes*, *11*(6), 648. https://doi.org/10.3390/genes11060648

Treff, N. R., Zimmerman, R., Bechor, E., Hsu, J., Rana, B., Jensen, J., Li, J., Samoilenko, A., Mowrey, W., Van Alstine, J., Leondires, M., Miller, K., Paganetti, E., Lello, L., Avery, S., Hsu, S., & Melchior Tellier, L. C. A. (2019). Validation of concurrent preimplantation genetic testing for polygenic and monogenic disorders, structural rearrangements, and whole and segmental chromosome aneuploidy with a single universal platform. *European Journal of Medical Genetics*, *62*(8), 103647. https://doi.org/10.1016/j.ejmg.2019.04.004

Visscher, P. M., & Wray, N. R. (2015). Concepts and Misconceptions about the Polygenic Additive Model Applied to Disease. *Human Heredity*, *80*(4), 165–170. https://doi.org/10.1159/000446931

Visscher, P. M., Wray, N. R., Zhang, Q., Sklar, P., McCarthy, M. I., Brown, M. A., & Yang, J. (2017). 10 Years of GWAS Discovery: Biology, Function, and Translation. *The American Journal of Human Genetics*, *101*(1), 5–22. https://doi.org/10.1016/j.ajhg.2017.06.005

Wald, N. J., & Old, R. (2019). The illusion of polygenic disease risk prediction. *Genetics in Medicine*, *21*(8), 1705–1707. https://doi.org/10.1038/s41436-018-0418-5

Watanabe, K., Stringer, S., Frei, O., Umićević Mirkov, M., de Leeuw, C., Polderman, T. J. C., van der Sluis, S., Andreassen, O. A., Neale, B. M., & Posthuma, D. (2019). A global overview of pleiotropy and genetic architecture in complex traits. *Nature Genetics*, *51*(9), 1339–1348. https://doi.org/10.1038/s41588-019-0481-0

Wilkinson, J., Malpas, P., Hammarberg, K., Mahoney Tsigdinos, P., Lensen, S., Jackson, E., Harper, J., & Mol, B. W. (2019). Do à la carte menus serve infertility patients? The ethics and regulation of in vitro fertility add-ons. *Fertility and Sterility*, *112*(6), 973–977. https://doi.org/10.1016/j.fertnstert.2019.09.028

Wray, N. R., & Goddard, M. E. (2010). Multi-locus models of genetic risk of disease. *Genome Medicine*, *2*(2), 10. https://doi.org/10.1186/gm131

Wray, N. R., Lin, T., Austin, J., McGrath, J. J., Hickie, I. B., Murray, G. K., & Visscher, P. M. (2020). From Basic Science to Clinical Application of Polygenic Risk Scores: A Primer. *JAMA Psychiatry*. https://doi.org/10.1001/jamapsychiatry.2020.3049

Wray, N. R., Yang, J., Hayes, B. J., Price, A. L., Goddard, M. E., & Visscher, P. M. (2013). Pitfalls of predicting complex traits from SNPs. *Nature Reviews. Genetics*, *14*(7), 507–515. https://doi.org/10.1038/nrg3457

Zhang, Q., Sidorenko, J., Couvy-Duchesne, B., Marioni, R. E., Wright, M. J., Goate, A. M., Marcora, E., Huang, K., Porter, T., Laws, S. M., Sachdev, P. S., Mather, K. A., Armstrong, N. J., Thalamuthu, A., Brodaty, H., Yengo, L., Yang, J., Wray, N. R., McRae, A. F., & Visscher, P. M. (2020). Risk

prediction of late-onset Alzheimer's disease implies an oligogenic architecture. *Nature Communications*, *11*(1), 4799. https://doi.org/10.1038/s41467-020-18534-1

Zhang, X., Rice, M., Tworoger, S. S., Rosner, B. A., Eliassen, A. H., Tamimi, R. M., Joshi, A. D., Lindstrom, S., Qian, J., Colditz, G. A., Willett, W. C., Kraft, P., & Hankinson, S. E. (2018). Addition of a polygenic risk score, mammographic density, and endogenous hormones to existing breast cancer risk prediction models: A nested case-control study. *PLoS Medicine*, *15*(9), e1002644. https://doi.org/10.1371/journal.pmed.1002644

Zhang, Y., Qi, G., Park, J.-H., & Chatterjee, N. (2018). Estimation of complex effect-size distributions using summary-level statistics from genome-wide association studies across 32 complex traits. *Nature Genetics*, *50*(9), 1318–1326. https://doi.org/10.1038/s41588-018-0193-x

Zheng, J., Erzurumluoglu, A. M., Elsworth, B. L., Kemp, J. P., Howe, L., Haycock, P. C., Hemani, G., Tansey, K., Laurin, C., Early Genetics and Lifecourse Epidemiology (EAGLE) Eczema Consortium, Pourcain, B. S., Warrington, N. M., Finucane, H. K., Price, A. L., Bulik-Sullivan, B. K., Anttila, V., Paternoster, L., Gaunt, T. R., Evans, D. M., & Neale, B. M. (2017). LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics (Oxford, England)*, *33*(2), 272–279. https://doi.org/10.1093/bioinformatics/btw613

Utility of polygenic embryo screening for disease depends on the selection strategy

Supplementary Methods

November 4, 2020

1 The liability threshold model

The liability threshold model (LTM) is a classic model in quantitative genetics [1, 2, 3], and is also commonly used to analyze modern data (e.g., [4, 5, 6, 7, 8, 9, 10, 11]). The LTM assumes that a disease has an underlying "liability", which is normally distributed in the population, and is the sum of two components: genetic and non-genetic (the environment). Further, the LTM assumes an "infinitesimal", or "polygenic" genetic basis, under which a very large number of genetic variants of small effect combine to form the genetic component. An individual is affected if his/her total liability (genetic + environmental) exceeds a threshold.

Mathematically, if we denote the liability as y, the LTM can be written as

$$y = g + \epsilon, \tag{1}$$

where $y \sim N(0, 1)$ is a standard normal variable, $g \sim N(0, h^2)$ is the genetic component, with variance equal to the (narrow-sense) heritability h^2, and $\epsilon \sim N(0, 1 - h^2)$ is the non-genetic component. In practice, we cannot measure the genetic component, but only estimate it imprecisely with a polygenic risk score, denoted s. Following previous work [5, 7, 12, 13], we assume that the LTM can be written, similarly to Eq. (1), as

$$y = s + e, \tag{2}$$

where $y \sim N(0, 1)$ as above, $s \sim N(0, r^2_{\text{ps}})$, where r^2_{ps} is the proportion of the variance in liability explained by the score, and $e \sim N(0, 1 - r^2_{\text{ps}})$ is the residual of the regression of the liability on s (and is uncorrelated with s), representing environmental effects as well as genetic factors not accounted for by the score.

An individual is affected whenever his/her liability exceeds a threshold, i.e., $y > \tau$. The threshold is selected such that the proportion of affected individuals

is equal to the prevalence K, i.e., $\tau = z_K$, where z_K is the $(1-K)$-quantile of a standard normal variable. Thus,

$$P(\text{disease}) = P(y > z_K) = K. \tag{3}$$

The model is illustrated in Figure 1.

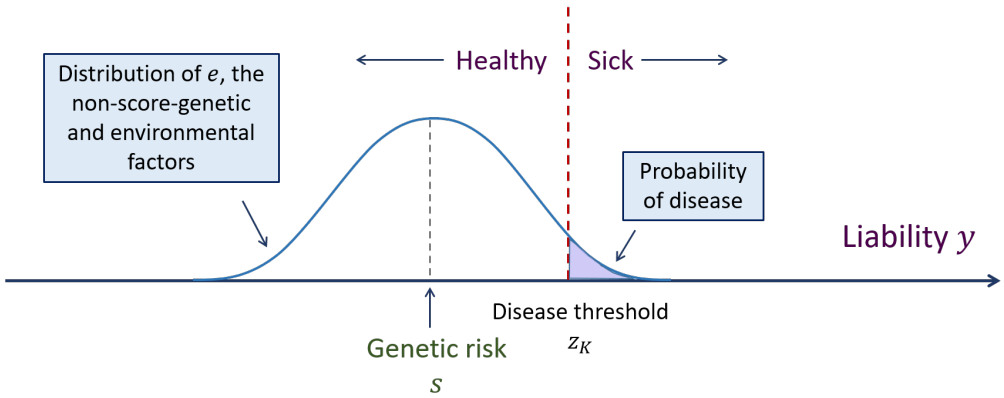


Figure 1: A schematic illustration of the liability threshold model. The liability is $y \sim s + e$, where s represents the polygenic score that quantifies the genetic risk (with variance r^2_{ps}), and e represents environmental factors as well as genetic factors that are not accounted for by the score (variance $1 - r^2_{\text{ps}}$). An individual is affected whenever the liability $y > z_K$, where z_K is the $(1-K)$-quantile of the standard normal distribution.

2 A model for the scores of n IVF embryos

Consider the polygenic risk scores (for a disease of interest) of n IVF embryos of given parents. We assume no information is known about the parents, or, in other words, that the parents are randomly drawn from the population. The scores of the embryos have a multivariate normal distribution,

$$\mathbf{s} = (s_1, \ldots, s_n) = \text{MVN}(\mathbf{0}_n, \mathbf{\Sigma}), \tag{4}$$

where the means form a vector of n zeros, and the $n \times n$ covariance matrix is

$$\mathbf{\Sigma} = r^2_{\text{ps}} \begin{pmatrix} 1 & \frac{1}{2} & \ldots & \frac{1}{2} \\ \frac{1}{2} & 1 & \ldots & \frac{1}{2} \\ \ldots & \ldots & \ldots & \ldots \\ \frac{1}{2} & \frac{1}{2} & \ldots & 1 \end{pmatrix}. \tag{5}$$

The diagonal elements of the matrix are simply the variances of the individual scores of each embryo. The off-diagonal elements represent the covariance between the scores of the embryos, who are genetically siblings. Based on standard

quantitative genetic theory [3] (see also our previous paper [13]), the covariance between the scores of two siblings is $\mathrm{Cov}\,(s_i, s_j) = \frac{1}{2}\mathrm{Var}\,(s)$, and hence the off-diagonal elements follow. [The non-score components (the e terms in Eq. (2)) are also correlated, as they represent both genetic and environmental factors. However, this does not affect the results of the present paper. See also Section 8 below.]

As we showed in our previous work [13], the scores can be written as a sum of two independent multivariate normal variables, $\mathbf{s} = \boldsymbol{x} + \boldsymbol{c}$, with

$$\begin{aligned} \boldsymbol{x} &= (x_1, \ldots, x_n) \sim \mathrm{MVN}\left(\mathbf{0}_n, \frac{r_{\mathrm{ps}}^2}{2}\boldsymbol{I}_n\right), \\ \boldsymbol{c} &= (c_1, \ldots, c_n) \sim \mathrm{MVN}\left(\mathbf{0}_n, \frac{r_{\mathrm{ps}}^2}{2}\boldsymbol{J}_n\right), \end{aligned} \tag{6}$$

where $\mathbf{0}_n$ is a vector of zeros of length n, $\boldsymbol{I}_n$ is the $n \times n$ identity matrix, and $\boldsymbol{J}_n$ is the $n \times n$ matrix of all ones. The x_i's and c_i's have the same marginal distribution, namely normal with mean zero and variance $r_{\mathrm{ps}}^2/2$ each. However, the x_i's are independent, whereas $\boldsymbol{c}$ has a constant covariance matrix, which means that the c_i's are n identical copies of the same random variable,

$$c_1 \sim N\left(0, \frac{r_{\mathrm{ps}}^2}{2}\right) \text{ and } c_2 = c_3 = \cdots = c_n = c_1 \equiv c. \tag{7}$$

Thus, for each embryo $i = 1, \ldots, n$,

$$s_i = x_i + c. \tag{8}$$

We note that in Eq. (8), c can also be interpreted as the mean polygenic score of the parents. In other words, given the mean scores of the parents (i.e., conditional on c), the scores of all siblings are independently and normally distributed, with mean c and variance $r_{\mathrm{ps}}^2/2$.

3 The disease risk when implanting the embryo with the lowest risk

We assume next that we select for implantation the embryo with the lowest polygenic risk score for the disease of interest. Our goal will be to calculate the probability of that embryo to be affected. Since $s_i = x_i + c$, the score of the selected embryo satisfies

$$\begin{aligned} s_{\min} &= \min(x_1 + c, \ldots, x_n + c) \\ &= \min(x_1, \ldots, x_n) + c \\ &= x_{\min} + c, \end{aligned} \tag{9}$$

where we defined $x_{\min} = \min(x_1, \ldots, x_n)$. Denote by i^* the index of the selected embryo ($x_{i^*} = x_{\min}$). The liability of the embryo with the lowest risk is thus

$$\begin{aligned} y_{i^*} &= s_{\min} + e_{i^*} \\ &= x_{\min} + c + e_{i^*} \\ &= x_{\min} + \tilde{e}, \end{aligned} \tag{10}$$

where $\tilde{e} = c + e_{i^*}$. Consequently,

$$\mathrm{Var}\,(\tilde{e}) = \mathrm{Var}\,(c) + \mathrm{Var}\,(e_{i^*}) = \frac{r_{\mathrm{ps}}^2}{2} + (1 - r_{\mathrm{ps}}^2) = 1 - \frac{r_{\mathrm{ps}}^2}{2}. \tag{11}$$

Therefore, the liability of the selected embryo can be written as a sum of two (independent) variables: $x_{\min}$, which is the minimum of n *independent* (zero mean) normal variables with variance $r_{\mathrm{ps}}^2/2$ each; and $\tilde{e}$, which is a normal variable with (zero mean and) variance $1 - r_{\mathrm{ps}}^2/2$.

The distribution of $x_{\min}$ can be computed based on the theory of order statistics,

$$P(x_{\min} > x_m) = [P(x > x_m)]^n = \left[1 - \Phi\left(\frac{x_m}{r_{\mathrm{ps}}/\sqrt{2}}\right)\right]^n. \tag{12}$$

In the above equation, the probability that the minimum of n variables is greater than x_m is the probability that all variables are greater than x_m. The distribution of each x is normal with zero mean and variance $r_{\mathrm{ps}}^2/2$, and hence $P(x > x_m) = 1 - \Phi\left(\frac{x_m}{r_{\mathrm{ps}}/\sqrt{2}}\right)$, where $\Phi(\cdot)$ is the cumulative probability distribution (CDF) of a standard normal variable.

We can now compute the probability of the selected embryo to be affected by the disease by demanding that the total liability is greater than the threshold z_K. Conditional on $\tilde{e}$,

$$\begin{aligned} P(\text{disease}\,|\,\tilde{e}) &= P(y_{i^*} > z_K\,|\,\tilde{e}) \\ &= P(x_{\min} + \tilde{e} > z_K) \\ &= P(x_{\min} > z_K - \tilde{e}) \\ &= \left[1 - \Phi\left(\frac{z_K - \tilde{e}}{r_{\mathrm{ps}}/\sqrt{2}}\right)\right]^n, \end{aligned} \tag{13}$$

where in the fourth line we used Eq. (12). Next, denote by $f(\tilde{e})$ the density of $\tilde{e}$, and by $\phi(\cdot)$ for the probability density function of the standard normal

variable. Given that $\tilde{e} \sim N(0, 1 - r_{\rm ps}^2/2)$,

$$
\begin{aligned}
P(\text{disease}) &= \int_{-\infty}^{\infty} P(\text{disease} \,|\, \tilde{e}) f(\tilde{e}) d\tilde{e} \\
&= \int_{-\infty}^{\infty} \left[1 - \Phi\left(\frac{z_K - \tilde{e}}{r_{\rm ps}/\sqrt{2}} \right) \right]^n \frac{1}{\sqrt{1 - r_{\rm ps}^2/2}} \phi\left(\frac{\tilde{e}}{\sqrt{1 - r_{\rm ps}^2/2}} \right) d\tilde{e} \\
&= \int_{-\infty}^{\infty} \left[1 - \Phi\left(\frac{z_K - t\sqrt{1 - r_{\rm ps}^2/2}}{r_{\rm ps}/\sqrt{2}} \right) \right]^n \phi(t) dt. \qquad (14)
\end{aligned}
$$

In the third line, we changed variables: $t = \tilde{e}/\sqrt{1 - r_{\rm ps}^2/2}$. Eq. (14) is our final expression for the probability of the embryo with the lowest score to be affected.

3.1 The risk reduction when conditioning on the mean parental score

Consider the case when c is given, or, in other words, that we know the mean parental polygenic score. Let us compute the disease risk in such a case. We start from Eq. (10),

$$
\begin{aligned}
y_{i^*} &= s_{\min} + e_{i^*} \\
&= x_{\min} + c + e_{i^*}. \qquad (15)
\end{aligned}
$$

Then,

$$
\begin{aligned}
P(\text{disease} \,|\, c, e_{i^*}) &= P(y_{i^*} > z_K \,|\, e_{i^*}) \\
&= P(x_{\min} + c + e_{i^*} > z_K) \\
&= P(x_{\min} > z_K - c - e_{i^*}) \\
&= \left[1 - \Phi\left(\frac{z_K - c - e_{i^*}}{r_{\rm ps}/\sqrt{2}} \right) \right]^n, \qquad (16)
\end{aligned}
$$

where in the last line we used Eq. (12).

Finally, recalling that $e_{i^*} \sim N\left(0, 1 - r_{\rm ps}^2\right)$,

$$
\begin{aligned}
P(\text{disease} \,|\, c) &= \int_{-\infty}^{\infty} P(\text{disease} \,|\, c, e_{i^*}) f(e_{i^*}) de_{i^*} \\
&= \int_{-\infty}^{\infty} \left[1 - \Phi\left(\frac{z_K - c - e_{i^*}}{r_{\rm ps}/\sqrt{2}} \right) \right]^n \frac{1}{\sqrt{1 - r_{\rm ps}^2}} \phi\left(\frac{e_{i^*}}{\sqrt{1 - r_{\rm ps}^2}} \right) de_{i^*} \\
&= \int_{-\infty}^{\infty} \left[1 - \Phi\left(\frac{z_K - c - t\sqrt{1 - r_{\rm ps}^2}}{r_{\rm ps}/\sqrt{2}} \right) \right]^n \phi(t) dt, \qquad (17)
\end{aligned}
$$

where in the last line we changed variables, $t = e_{i^*}/\sqrt{1 - r_{\text{ps}}^2}$. Eq. (17) thus provides the probability of disease when we are given the mean parental score c.

4 The disease risk when excluding high-risk embryos

We now consider the selection strategy in which the embryo for implantation is selected at random, as long as its risk score is not particularly high. Specifically, we assume that whenever possible, embryos at the top q risk percentiles are excluded. When *all* embryos have high risk, we assume that a random embryo is selected. Let z_q be the $(1-q)$-quantile of the standard normal distribution. The variance of the score is r_{ps}^2, and therefore, the score of the selected embryo must be lower than $z_q r_{\text{ps}}$.

To compute the disease risk in this case, we first condition on the shared, family-specific component c. We later integrate over c to derive the risk across the population. Denote by x_s the value of x for the *selected* embryo, and for the moment, also condition on x_s. We have,

$$\begin{aligned} P(\text{disease} \,|\, x_s, c) &= P(y > z_K \,|\, c) \\ &= P(s + e > z_K) \\ &= P(x_s + c + e > z_K) \\ &= P(e > z_K - x_s - c) \\ &= 1 - \Phi\left(\frac{z_K - x_s - c}{\sqrt{1 - r_{\text{ps}}^2}}\right), \end{aligned} \tag{18}$$

To obtain $P(\text{disease} \,|\, c)$, we need to integrate over $f(x_s)$, the density of x_s. In fact, $f(x_s)$ is a mixture of two distributions, depending on whether or not all embryos were high risk. Denote by H the event that all embryos are high risk, and let us first compute the probability of H. Recall that given c, the scores of all embryos, $s_i = x_i + c$, are independent. The event H is equivalent to the intersection of the *independent* events $s_i > z_q r_{\text{ps}}$ for $i = 1, \ldots, n$. Thus,

recalling that $x_i \sim N(0, r_{\text{ps}}^2/2)$,

$$\begin{aligned} P(H) &= \prod_{i=1}^{n} P(s_i > z_q r_{\text{ps}}) \\ &= \prod_{i=1}^{n} P(x_i + c > z_q r_{\text{ps}}) \\ &= \prod_{i=1}^{n} P(x_i > z_q r_{\text{ps}} - c) \\ &= \left[1 - \Phi\left(\frac{z_q r_{\text{ps}} - c}{r_{\text{ps}}/\sqrt{2}}\right)\right]^n . \end{aligned} \tag{19}$$

Given H, we know that all scores were higher than the cutoff, i.e., that $x_i > z_q r_{\text{ps}} - c$ for all $i = 1, \ldots, n$. An embryo is then selected at random, and again, we recall that the x_i's are independent. Thus, x_s, the value of x of the selected embryo, is a realization of a normal random variable truncated from below. Specifically, if $f_x(\cdot)$ is the unconditional density of x, then for $x_s > z_q r_{\text{ps}} - c$,

$$f(x_s \,|\, H) = \frac{f_x(x_s)}{P(x > z_q r_{\text{ps}})} = \frac{\frac{1}{r_{\text{ps}}/\sqrt{2}}\phi\left(\frac{x_s}{r_{\text{ps}}/\sqrt{2}}\right)}{1 - \Phi\left(\frac{z_q r_{\text{ps}} - c}{r_{\text{ps}}/\sqrt{2}}\right)}. \tag{20}$$

In the case H did not occur, we select an embryo at random among embryos with score $s_i < z_q r_{\text{ps}}$, i.e., $x_i < z_q r_{\text{ps}} - c$. The density of x_s is again, analogously to the above case, a realization of a normal random variable, but this time truncated from above. For $x_s < z_q r_{\text{ps}} - c$,

$$f(x_s \,|\, \overline{H}) = \frac{f_x(x_s)}{P(x < z_q r_{\text{ps}})} = \frac{\frac{1}{r_{\text{ps}}/\sqrt{2}}\phi\left(\frac{x_s}{r_{\text{ps}}/\sqrt{2}}\right)}{\Phi\left(\frac{z_q r_{\text{ps}} - c}{r_{\text{ps}}/\sqrt{2}}\right)}. \tag{21}$$

Using these results, we can write the unconditional density of x_s,

$$\begin{aligned} f(x_s) &= \begin{cases} f(x_s \,|\, H)P(H) + 0 \cdot P(\overline{H}) & \text{for } x_s > z_q r_{\text{ps}} - c \\ 0 \cdot P(H) + f(x_s \,|\, \overline{H})P(\overline{H}) & \text{for } x_s < z_q r_{\text{ps}} - c \end{cases} \\ &= \begin{cases} \frac{\frac{1}{r_{\text{ps}}/\sqrt{2}}\phi\left(\frac{x_s}{r_{\text{ps}}/\sqrt{2}}\right)}{1-\Phi\left(\frac{z_q r_{\text{ps}} - c}{r_{\text{ps}}/\sqrt{2}}\right)} \left[1 - \Phi\left(\frac{z_q r_{\text{ps}} - c}{r_{\text{ps}}/\sqrt{2}}\right)\right]^n & \text{for } x_s > z_q r_{\text{ps}} - c \\ \frac{\frac{1}{r_{\text{ps}}/\sqrt{2}}\phi\left(\frac{x_s}{r_{\text{ps}}/\sqrt{2}}\right)}{\Phi\left(\frac{z_q r_{\text{ps}} - c}{r_{\text{ps}}/\sqrt{2}}\right)} \left\{1 - \left[1 - \Phi\left(\frac{z_q r_{\text{ps}} - c}{r_{\text{ps}}/\sqrt{2}}\right)\right]^n\right\} & \text{for } x_s < z_q r_{\text{ps}} - c \end{cases} \end{aligned} \tag{22}$$

We can now integrate over all x_s, still conditioning on c, and using Eq. (18) and some algebra,

$$
\begin{aligned}
P(\text{disease}\,|\,c) &= \int_{-\infty}^{\infty} f(x_s) P(\text{disease}\,|\,x_s, c) dx_s \\
&= \int_{-\infty}^{z_q r_{\text{ps}} - c} \frac{\frac{1}{r_{\text{ps}}/\sqrt{2}} \phi\left(\frac{x_s}{r_{\text{ps}}/\sqrt{2}}\right)}{\Phi\left(\frac{z_q r_{\text{ps}} - c}{r_{\text{ps}}/\sqrt{2}}\right)} \left\{1 - \left[1 - \Phi\left(\frac{z_q r_{\text{ps}} - c}{r_{\text{ps}}/\sqrt{2}}\right)\right]^n\right\} \left[1 - \Phi\left(\frac{z_K - x_s - c}{\sqrt{1 - r_{\text{ps}}^2}}\right)\right] dx_s \\
&+ \int_{z_q r_{\text{ps}} - c}^{\infty} \frac{\frac{1}{r_{\text{ps}}/\sqrt{2}} \phi\left(\frac{x_s}{r_{\text{ps}}/\sqrt{2}}\right)}{1 - \Phi\left(\frac{z_q r_{\text{ps}} - c}{r_{\text{ps}}/\sqrt{2}}\right)} \left[1 - \Phi\left(\frac{z_q r_{\text{ps}} - c}{r_{\text{ps}}/\sqrt{2}}\right)\right]^n \left[1 - \Phi\left(\frac{z_K - x_s - c}{\sqrt{1 - r_{\text{ps}}^2}}\right)\right] dx_s \\
&= \int_{-\infty}^{\infty} \eta(t) \xi(t) dt, \qquad (23)
\end{aligned}
$$

where we defined

$$
\begin{aligned}
\gamma &= \sqrt{2} z_q - \frac{c}{r_{\text{ps}}/\sqrt{2}}, \\
\xi(t) &= \phi(t) \left[1 - \Phi\left(\frac{z_K - t r_{\text{ps}}/\sqrt{2} - c}{\sqrt{1 - r_{\text{ps}}^2}}\right)\right], \text{ and} \\
\eta(t) &= \begin{cases} \frac{1 - [1 - \Phi(\gamma)]^n}{\Phi(\gamma)} & \text{for } t < \gamma, \\ \frac{[1 - \Phi(\gamma)]^n}{1 - \Phi(\gamma)} & \text{for } t > \gamma. \end{cases} \qquad (24)
\end{aligned}
$$

Eq. (23) provides an expression for the probability of disease given the mean parental score c.

Finally, we can integrate over all c in order to obtain the probability of disease in the population. Recalling that $c \sim N(0, r_{\text{ps}}^2/2)$ and denoting its density as $f(c)$, and again after some algebra,

$$
\begin{aligned}
P(\text{disease}) &= \int_{-\infty}^{\infty} P(\text{disease}\,|\,c) f(c) dc \\
&= \int_{-\infty}^{\infty} \phi(u) \left[\int_{-\infty}^{\infty} \psi(u,t) \zeta(u,t) dt\right] du, \qquad (25)
\end{aligned}
$$

where we defined

$$
\begin{aligned}
\beta &= \sqrt{2} z_q - u, \\
\zeta(u,t) &= \phi(t) \left[1 - \Phi\left(\frac{z_K - (u+t) r_{\text{ps}}/\sqrt{2}}{\sqrt{1 - r_{\text{ps}}^2}}\right)\right], \text{ and} \\
\psi(u,t) &= \begin{cases} \frac{1 - [1 - \Phi(\beta)]^n}{\Phi(\beta)} & \text{for } t < \beta, \\ \frac{[1 - \Phi(\beta)]^n}{1 - \Phi(\beta)} & \text{for } t > \beta. \end{cases} \qquad (26)
\end{aligned}
$$

Eq. (25) is our final expression for the probability of an embryo selected randomly among non-high-risk embryos to be affected.

5 The relative risk reduction

Given the prevalence K and the probability of the selected embryo to be affected $P(\text{disease})$, we define the relative risk reduction (RRR) as

$$\text{RRR} = \frac{K - P(\text{disease})}{K} = 1 - \frac{P(\text{disease})}{K}. \tag{27}$$

For example, if a disease has prevalence of 5% and the selected embryo has probability 3% to be affected, the relative risk reduction is 40%. We use the relative risk reduction defined in Eq. (27) in the main text, where $P(\text{disease})$ is given by Eq. (14) for the *lowest-risk prioritization* strategy, and by Eq. (25) for the *high-risk exclusion* strategy. We solve the integrals in these equations numerically in R (see Section 9).

5.1 The relative risk reduction conditional on the mean parental score

In Sections 3 and 4 we computed the probability of disease conditional on the mean parental score c for the two selection strategies (Eqs. (17) and (23)). Let us denote this probability as $P_s(\text{disease}\,|\,c)$. To compute the relative risk reduction in these cases, we also need the "baseline" risk, i.e., the probability of disease when a random embryo is selected (as in natural procreation). We denote this probability $P_r(\text{disease}\,|\,c)$. Let us write the liability of a random embryo as

$$\begin{aligned} y &= s + e \\ &= x + c + e \\ &= \tilde{x} + c, \end{aligned} \tag{28}$$

where we defined $\tilde{x} = x + e$. $\text{Var}\,(\tilde{x}) = \text{Var}\,(x) + \text{Var}\,(e) = r_{\text{ps}}^2/2 + 1 - r_{\text{ps}}^2 = 1 - r_{\text{ps}}^2/2$, and thus, $\tilde{x} \sim N\left(0, 1 - r_{\text{ps}}^2/2\right)$. The conditional probability of disease is

$$\begin{aligned} P_r(\text{disease}\,|\,c) &= P(y > z_K\,|\,c) \\ &= P(\tilde{x} + c > z_K) \\ &= P(\tilde{x} > z_K - c) \\ &= 1 - \Phi\left(\frac{z_K - c}{\sqrt{1 - r_{\text{ps}}^2/2}}\right). \end{aligned} \tag{29}$$

The relative risk reduction conditional on the mean parental score can be computed as

$$\text{RRR} = \frac{P_r(\text{disease}\,|\,c) - P_s(\text{disease}\,|\,c)}{P_r(\text{disease}\,|\,c)} = 1 - \frac{P_s(\text{disease}\,|\,c)}{P_r(\text{disease}\,|\,c)}. \tag{30}$$

The absolute risk reduction is defined as $P_r(\text{disease}\,|\,c) - P_s(\text{disease}\,|\,c)$.

6 Comparison to previous work

In the "gwern" blog post [14], the utility of embryo selection for traits and/or diseases was investigated. For disease risk, a similar model to ours was studied, based on the liability threshold model. However, the model assumed that given the polygenic score, the distribution of the remaining contribution to the liability has unit variance, instead of $1 - r_{\mathrm{ps}}^2$ (the function `liabilityThresholdValue` therein). Moreover, only numerical results were provided, the *high-risk exclusion* strategy was not considered, and there was no analysis of the case when the parental scores are given. Treff et al. [12] also employed the liability threshold model to evaluate embryo selection for disease risk. However, they did not consider the *high-risk exclusion* strategy, and did not compute analytically the risk reduction. They only provided simulation results for the case when a parent is affected based on an approximate model.

7 Simulations

To simulate the outcomes of embryo selection, we use the representation $s_i = x_i + c$, where $(x_1, \ldots, x_n)$ are independent normals with zero means and variance $r_{\mathrm{ps}}^2/2$, and $c \sim N(0, r_{\mathrm{ps}}^2/2)$ is shared across all embryos. Thus, for each "family", we first draw c, then draw n independent normals $(x_1, \ldots, x_n)$, and then compute the score of embryo i as $s_i = x_i + c$, for $i = 1, \ldots, n$. The score of the selected embryo is the lowest among the n embryos in the *lowest-risk prioritization* strategy. For the *high-risk exclusion* strategy, we select the first embryo with score $s < z_q r_{\mathrm{ps}}$. If no such embryo exists, we select the first embryo. We then draw the residual of the liability as $e \sim N(0, 1 - r_{\mathrm{ps}}^2)$, and compute the liability as $s^* + e$, where s^* is the score of the selected embryo. If the liability exceeds the threshold z_K, we designate the embryo as affected. We repeat over 10^6 sets of embryos, and compute the disease probability as the fraction of "families" in which the selected embryo was affected. We compute the relative risk reduction using Eq. (27).

8 Limitations of the model

Our model has a number of limitations. First, we assumed an infinitesimal genetic architecture for the disease, which will not be appropriate for oligogenic diseases or when screening the embryos for variants of very large effect. Our model also implicitly assumes no assortative mating, which seems reasonable given that for genetic disease risk, correlation between parents is weak [15], and given that our previous study of traits showed no difference in the results between real and random couples [13].

Second, we assumed that the proportion of variance (on the liability scale) explained by the score is r_{ps}^2, but we did not specify how to estimate it. Typically, r_{ps}^2 is computed and reported by large GWASs based on an evaluation of the score in a test set. However, the variance that will be explained by the score in

other cohorts, and particularly, in other populations, can be substantially lower [16]. Relatedly, the variance explained by the score, as estimated in samples of unrelated individuals, is inflated due to population stratification, assortative mating, and indirect parental effects ("genetic nurture") [17, 18, 19, 20]. These effects are absent when comparing polygenic scores between siblings (as when screening IVF embryos), and thus, the variance explained by polygenic scores in this setting can be substantially reduced, in particular for cognitive traits. However, recent empirical work on within-family disease risk prediction showed that the reduction in accuracy is at most modest [21].

Third, we do not model the process of IVF and the possible reasons for loss of embryos. Rather, we assume that n viable embryos are available that would have led to live birth if implanted. The original number of fertilized oocytes would typically be greater than n (see, e.g., [14] for more detailed modeling). Similarly, we do not model the age-dependence of the number of embryos; again, we rather assume n viable embryos are available.

Fourth, the residual e in Eq. (2) ($y = s + e$) has a complex pattern of correlation between siblings. As noted in Section 2, e has contribution from both genetic and environmental factors. While the genetic covariance between siblings is straightforward to model (as in Section 2), the proportion of variance in liability explained by shared environment needs to be estimated and can be large [22]. Further, embryos from the same IVF cycle (when only one is actually implanted) would have experienced the same early developmental environment, and are thus expected to share even more environmental factors, similarly to twins. In the current work, the correlation between the residuals across embryos does not enter our derivations. However, care must be taken in any future attempt to compare the predicted phenotypic outcomes across the embryos.

Finally, in this work we only modeled the case when the parents are either randomly selected from the population or when their scores are given, and when screening the embryos for a single disease. In future work, it will be interesting to model the case when one or more parents are affected or when the parents already have an affected child. Further, it is likely that parents would attempt to screen the embryos for more than one disease [23]. A related important question for future studies is to what extent selecting for a lower risk of one disease (or a set of diseases) would increase risk for diseases that were not included in the screen.

9 Code availability

The R code we used to compute the relative risk reduction for the *lowest-risk prioritization* strategy is as follows.

```
library(MASS)
risk_reduction_lowest = function(r,K,n)
{
  zk = qnorm(K, lower.tail=F)
  integrand_lowest = function(t)
```

```
    return(dnorm(t)*pnorm((zk-t*sqrt(1-r^2/2)) / (r/sqrt(2)), lower.tail=F)^n)
  risk = integrate(integrand_lowest,-Inf,Inf)$value
  return((K-risk)/K)
}
```

The R code we used to compute the relative risk reduction for the *high-risk exclusion* strategy is as follows.

```
risk_reduction_exclude = function(r,K,q,n)
{
  zk = qnorm(K, lower.tail=F)
  zq = qnorm(q, lower.tail=F)
  integrand_t = function(t,u)
    return(dnorm(t)*pnorm((zk-r/sqrt(2)*(u+t))/sqrt(1-r^2),lower.tail=F))
  integrand_u = function(us)
  {
    y = numeric(length(us))
    for (i in seq_along(us))
    {
      u = us[i]
      beta = zq*sqrt(2)-u
      internal_int1 = integrate(integrand_t,-Inf,beta,u)$value
      denom1 = pnorm(beta)
      if (denom1==0) {denom1=1e-300} # Avoid dividing by zero
      numer1 = 1-pnorm(beta,lower.tail=F)^n
      internal_int2 = integrate(integrand_t,beta,Inf,u)$value
      denom2 = pnorm(beta,lower.tail=F)
      if (denom2==0) {denom2=1e-300} # Avoid dividing by zero
      numer2 = pnorm(beta,lower.tail=F)^n
      y[i] = dnorm(u) * (numer1/denom1*internal_int1 + numer2/denom2*internal_int2)
    }
    return(y)
  }
  risk = integrate(integrand_u,-Inf,Inf)$value
  return((K-risk)/K)
}
```

The remaining code used to generate the figures of the main text can be found at `https://github.com/scarmi/embryo_selection`.

References

[1] E. R. Dempster and I. M. Lerner. Heritability of threshold characters. *Genetics*, 35:212, 1950.

[2] D. S. Falconer. The inheritance of liability to certain diseases, estimated from the incidence among relatives. *Ann Hum Genet*, 29:51, 1965.

[3] M. Lynch and B. Walsh. *Genetics and Analysis of Quantitative Traits*. Sinauer Associates, 1998.

[4] N. R. Wray and M. E. Goddard. Multi-locus models of genetic risk of disease. *Genome Med*, 2:10, 2010.

[5] H.-C. So, J. H. S. Kwan, S. S. Cherny, and P. C. Sham. Risk prediction of complex diseases from familyhistory and known susceptibility loci,with applications for cancer screening. *Am J Hum Genet*, 88:548, 2011.

[6] S. H. Lee, N. R. Wray, M. E. Goddard, , and P. M. Visscher. Estimating missing heritability for diseasefrom genome-wide association studies. *Am J Hum Genet*, 88:294, 2011.

[7] S. H. Lee, M. E. Goddard, N. R. Wray, and P. M. Visscher. A better coefficient of determination for genetic profile analysis. *Genet Epidemiol*, 36:214, 2012.

[8] C. B. Do, D. A. Hinds, U. Francke, and N. Eriksson. Comparison of family history and snps for predictingrisk of complex disease. *PLOS Genet*, 8:e1002973, 2012.

[9] T. J. Hayeck, P.-R. Loh, S. Pollack, A. Gusev, N. Patterson, Zaitlen N. A, and A. L. Price. Mixed model association with family-biased case-control ascertainment. *Am J Hum Genet*, 100:31, 2017.

[10] O. Weissbrod, J. Flint, and S. Rosset. Estimating snp-based heritability and genetic correlation in case-control studies directly and with summary statistics. *Am J Hum Genet*, 103:89, 2018.

[11] M. L. A. Hujoel, S. Gazal, P.-R. Loh, N. Patterson, and A. L. Price. Liability threshold modeling of case-control status and family history of disease increases association power. *Nat Genet*, 52:541, 2020.

[12] N. R. Treff, J. Eccles, L. Lello, E. Bechor, J. Hsu, K. Plunkett, R. Zimmerman, B. Rana, A. Samoilenko, S. Hsu, and L. C. A. M. Tellier. Utility and first clinical application of screening embryos for polygenic disease risk reduction. *Front Endocrinol*, 10:845, 2019.

[13] E. Karavani, O. Zuk, D. Zeevi, N. Barzilai, N. C. Stefanis, A. Hatzimanolis, N. Smyrnis, D. Avramopoulos, L. Kruglyak, G. Atzmon, M. Lam, T. Lencz, and S. Carmi. Screening human embryos for polygenic traits has limited utility. *Cell*, 179:1424, 2019.

[14] G. Branwen. Embryo selection for intelligence. *https://www.gwern.net/Embryo-selection*.

[15] K. Rawlik, O. Canela-Xandri, and A. Tenesa. Indirect assortative mating for human disease and longevity. *Heredity*, 123:106, 2019.

[16] A. R. Martin, M. Kanai, Y. Kamatani, Y. Okada, B. M. Neale, and M. J. Daly. Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat Genet*, 51:584, 2019.

[17] A. Kong, G. Thorleifsson, M. L. Frigge, B. J. Vilhjalmsson, A. I. Young, T. E. Thorgeirsson, S. Benonisdottir, A. Oddsson, B. V. Halldorsson, G. Masson, D. F. Gudbjartsson, A. Helgason, G. Bjornsdottir, U. Thorsteinsdottir, and K. Stefansson. The nature of nurture:effects of parental genotypes. *Science*, 359:424, 2018.

[18] A. I. Young, S. Benonisdottir, M. Przeworski, and A. Kong. Deconstructing the sources ofgenotype-phenotype associationsin humans. *Science*, 365:1396, 2019.

[19] T. T. Morris, N. M. Davies, G. Hemani, and G. Davey Smith. Population phenomena inflate genetic associationsof complex social traits. *Sci Adv*, 6:eaay0328, 2020.

[20] H. Mostafavi, A. Harpak, I. Agarwal, D. Conley, J. K Pritchard, and M. Przeworski. Variable prediction accuracy of polygenic scores within an ancestry group. *eLife*, 9:e48376, 2020.

[21] L. Lello, T. G. Raben, and S. D. H. Hsu. Sibling validation of polygenic riskscores and complex trait prediction. *Sci Rep*, 10:13190, 2020.

[22] C. M. Lakhani, B. T. Tierney, A. K. Manrai, J. Yang, P. M. Visscher, and C. J. Patel. Repurposing large health insurance claims data toestimate genetic and environmental contributionsin 560 phenotypes. *Nat Genet*, 51:327, 2019.

[23] N. R. Treff, J. Eccles, D. Marin, E. Messick, L. Lello, J. Gerber, J. Xu, and L. C. A. M. Tellier. Preimplantation genetic testing for polygenic disease relative risk reduction: Evaluation of genomic index performance in 11,883 adult sibling pairs. *Genes*, 11:648, 2020.