TANDEM MEETINGS

Transplantation & Cellular Therapy Meetings
of ASTCT® and CIBMTR®

SAMPLE SIZE AND CUT-OFF: IMPLICATIONS FOR INTRA-POPULATION MATCHING PROBABILITIES

J. Sauter, J. Pingel, A. H. Schmidt
DKMS Group, Tübingen, Germany

DKMS
WE DELETE BLOOD CANCER

INTRODUCTION

The probability to find a matching donor in a cohort of potential adult unrelated hematopoietic stem cell donors is a key determinant in registry operations. It guides decisions on donor centers' recruitment strategies, patients' therapy options as well as strategic registry planning[1] and research.

OBJECTIVES

While the matching probability monotonously increases and asymptotically converges toward 1 with registry size, its absolute value depends on the ethnic background of donor and patient as well as other parameters. Within our analyses, we explore the impact of different sample sizes and frequency cut-offs.

METHODS

The matching probability (MP) can be computed in a three-step process. Starting with a sample cohort of the population of interest, haplotype frequencies are estimated using established maximum-likelihood techniques[2,3,4]. Then, by assumption of Hardy-Weinberg Equilibrium, genotype frequencies are derived. In the third step, the MP is obtained via $MP = \sum g_i[1-(1-g_i)^n]$, where g_i is the genotype frequency and n the registry size.

RESULTS

For 10/10 matches, we systematically investigated the impact of sample sizes and frequency cut-offs for five exemplary populations based on donors with, self-assigned German (DE), Indian (IN), Polish (PL), Great Britain (UK), and non-indigenous Chilean (CL) ethnic background, respectively.

Investigated initial sample sizes include:

n=5,000, n=10,000, n=20,000, n=50,000, n=100,000.

Cut-off criteria for the low-frequency tail of the haplotype frequency distribution:

- At a haplotype frequency of 1/2n (corresponding to at least one occurrence of a haplotype in the initial sample)
- Cut-off at 99.5% of the cumulative sum of frequencies if sorted in descending order

Resulting haplotype frequency distributions (e.g. Figure 1) have been subsequently renormalized: $\sum f_i \stackrel{\text{def}}{=} 1$.

Figure 1

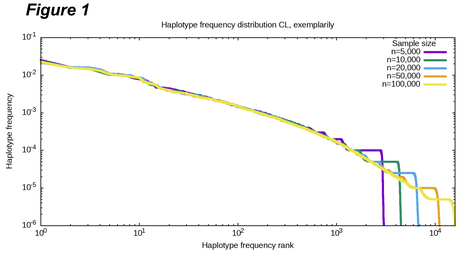


Figure 3

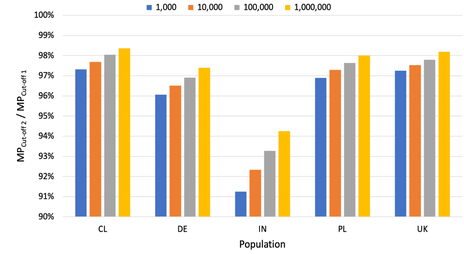
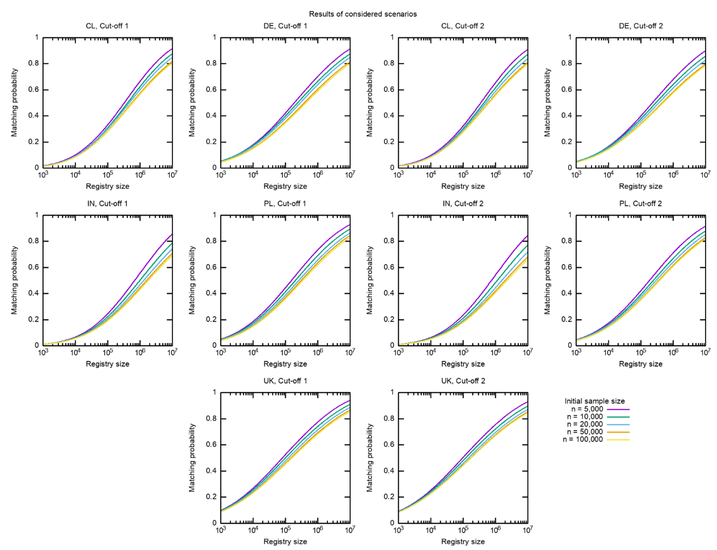


Figure 2



We computed the matching probability (MP) in all 50 scenarios (5 populations, 5 sample sizes, 2 cut-offs) for registry sizes ranging from 1,000 to 10,000,000 (Figure 2). For all populations and cut-offs, our results reveal for smaller registries a computed MP at similar values, irrespective of sample size. For larger registries, estimated MP is higher if based on smaller sample sizes. For example, with the 99.5% cut-off and a registry with one million donors, observed differences in MP estimation from the n=5,000 and n=100,000 sample range from 8.5% (UK) to 13.4% (IN) while for a registry with one thousand donors, result differences range from 0.1% (IN) to 0.7% (PL, UK). Concerning the two cut-offs, Figure 3 shows the quotient of the average MP from the 99.5 and 1/2n cut-off. Generally, MPs from the 1/2n cut-off yield between 92.2% (IN) and 98.3% (CL) of the MP from the other cut-off.

CONCLUSIONS

Our results can be understood considering the respectively included haplotypes as determined by parameter choices. Smaller samples and the 1/2n cut-off yield smaller numbers of haplotypes with the first also tending to neglect infrequent genotypes.

To further quantify the observed differences in terms of the samples' randomness, we plan to iterate over the respective sizes as an extension of the analyses.

REFERENCES

1 **Sauter J et al**. Optimisation of global stem cell donor recruitment based on analysis of unsuccessful donor searches. *HLA. 2024;104:e15610.*

2 **Schäfer C et al**. Hapl-o-Mat: open-source software for HLA haplotype frequency estimation from ambiguous and heterogeneous data. *BMC Bioinformatics. 2017/05/30 2017;18(1):284.*

3 **Solloch UV et al**. Graphical user interface for the haplotype frequency estimation software Hapl-o-Mat. *Human Immunology*. 2022/02/01/ 2022;83(2):107-112.

4 **`www.github.com/DKMS/Hapl-o-Mat`**

CONTACT INFORMATION

Author contact
sauter@dkms.de

DKMS
WE DELETE BLOOD CANCER