European Journal of Operational Research 000 (2016) 1–10

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Modeling the optimal ethnic composition of an adult stem cell registry

John T. Blake[a,b,*], Ken McTaggart[b], Donna Killeen[b]

[a] Dalhousie University, Department of Industrial Engineering, 5269 Morris Street, Halifax, NS B3H 4R2, Canada
[b] Canadian Blood Services, Medical Services and Innovation, Ottawa, Ontario, Canada

ARTICLE INFO

Article history:
Received 29 September 2015
Accepted 9 May 2016
Available online xxx

Keywords:
OR in health services
Stem cell registry composition
Large scale optimization
Heuristics

ABSTRACT

For patients suffering from a blood related disease, a stem cell transplant represents the best, and sometimes the only, possible treatment. Registries have been created throughout the world to match patients with stem cell donors. Canada's adult registry, OneMatch, was formed to meet the needs of Canadian patients. However, only 20–30 percent of unrelated adult stem cell transplants in Canada are sourced from Canadian donors. Self-sufficiency has proven difficult for OneMatch, in part, because the Canadian registry is dwarfed by other international registries with similar donor populations.

In this paper, we present a study to evaluate changes the Canadian registry designed to promote ethnic diversity while meeting the needs of the Canadian patient population. We formulate the composition problem as a linear optimization model and solve it using a combination of exact and heuristic methods for a registry of 1 M donors.

We conclude that when registry size is constrained, there are advantages to increasing ethnic diversity over self-sufficiency. However, results show that some communities cannot be easily accommodated within an adult registry of a fixed size. Thus, our results highlight the need for stem cells derived from cord blood for hard to match populations.

1. Introduction

For individuals suffering from leukemia or other blood related diseases, a stem cell transplant represents the best, and sometimes the only, possible course of treatment. While a genetically matched relative is the preferred donor source, in many cases a suitable relative cannot be found, and hence there is a need for a transplant from an unrelated individual. Throughout the world, cord blood banks and adult stem cell registries have been created to ensure patients in need can be matched with a suitable source of stem cells.

Stem cells for transplantation are typically obtained either from volunteer adult donors via apheresis collection methods or extracted from the blood of a newborn's umbilical cord and placenta. There are costs and benefits associated with both sources. Cord derived stem cells can be matched to a greater number of potential patients, since the matching criteria for cord stem cells is less stringent than for adults (Brunstein et al., 2010). In addition, because cords are collected at birth and frozen for later use they are always available for transplantation, whereas adult derived cells require that a potential donor, once identified, can be found and is willing to undergo the donation process (Sheehan-Connor, Bergstrom, & Garratt, 2015). Cord derived stem cells, however, are much more expensive than those from an adult source, have clinical limitations due to lower cell count than adult-derived cells, and are slower to engraft (Brunstein et al., 2010). In this paper we focus on the composition of Canada's adult stem cell registry, OneMatch[1], but assume the existence of cord derived stem cells to serve patients for whom no adult donor can be identified.

OneMatch, initiated in 1989 and containing information on 400,000 potential donors, was formed with the intent of meeting the needs of Canadian patients. However, less than 30 percent of unrelated adult stem cell transplants in Canada are sourced from Canadian donors; most Canadian patients receive transplants from international donors. This trend is not unique to Canada; approximately 45 percent of all adult stem cell transplants in the world involve a donor and a recipient from different countries (Sheehan-Connor, Bergstrom, & Garrat, 2015).

* Corresponding author at: Department of Industrial Engineering, Dalhousie University, 5269 Morris Street, Halifax, NS B3H 4R2, Canada. Tel.: +1 902 494 6068.
E-mail address: john.blake@dal.ca (J.T. Blake).

[1] Please note that OneMatch manages the stem cell registries in all provinces of Canada, except the Province of Québec which maintains a separate set of registries.

http://dx.doi.org/10.1016/j.ejor.2016.05.015
0377-2217/

There are several reasons why self-sufficiency has proven elusive for the OneMatch registry. The Canadian registry is small relative to other registries; the US National Marrow Donor Program has over 10.5 million registrants, while the German ZKRD program contains in excess of 6.0 million registrants. Furthermore, while Canada is an immigrant nation with an increasing ethnic diversity, the individuals recruited to the OneMatch program have historically been largely drawn from populations with European origins and thus there is a significant degree of overlap between the genetic makeup of Canadian registrants and those of the US and German programs. Additionally, the typical Canadian registrant, who is female, Caucasian, and over the age of 35, contrasts with current trends in donor selection by transplant centers (Bergstrom, Garratt, & Sheehan-Connor, 2009), in which younger male donors are preferred, because of their higher stem cell counts and the reduced risk of transplant rejection associated with males.

In 2014, Canadian Blood Services, which manages the OneMatch registry, undertook a review of its adult and cord blood (stem cells derived a newborn's umbilical cord and placenta) stem cell programs. This included benchmarking studies vis-à-vis other international registries, structured interviews with Canadian transplant centers, and models to evaluate the size (Blake, McTaggart, & Killeen, 2014) and composition of the Canadian registry. In this paper, we discuss the design, development, and use of a set of models to evaluate options for changing the demographic composition of the Canadian OneMatch registry in such a way as to promote ethnic diversity, while ensuring that the needs of the Canadian patient population can be met from either domestic or international sources.

2. Literature review

2.1. Stem cells: alleles, haplotypes, and phenotypes

Hematopoietic stem cell transplantation is a procedure used in the treatment of hematologic disorders, such as leukemia, lymphoma, and aplastic anemia (Sheehan-Connor et al., 2015). To ensure successful treatment and long term survival of the recipient, it is vital that the donor and the recipient be genetically matched at the human leukocyte antigen (HLA) level. The HLA system consists of more than 200 different genes, located on chromosome 6. Each person has a copy of chromosome 6 coming from their mother and their father; each copy is called a haplotype. Combined, the two haplotypes constitute a phenotype. The function of the HLA system is to signal to the immune system whether or not a cell is native (National Institute of Health, 2014).

HLA genes have many possible variations, which enable a person's immune system to react to a variety of foreign cells. There are three main classes of genes. Class I genes (HLA-A, HLA-B, and HLA-C) produce proteins that are present on the surface of all cells that are primary signals to the immune system of foreign versus native. Class II genes (HLA-DPB1, HLA-DQA1, HLA-DQB1, HLA-DRA, and HLA-DRB1) also provide instructions for making proteins that appear on the cell surface that regulate immune system response. Class III genes regulate inflammation and other immune system activities (National Institute of Health, 2014). In keeping with current clinical standards, the HLA-A, HLA-B, HLA-C, HLA-DRB1, and HLA-DQB1 genes are considered the most important for this study (Bergstrom et al., 2012). These genes are located at specific locations (loci) along the chromosome. Since there may be hundreds of known gene variants (alleles) associated with each locus in the HLA system, there are billions of potentially unique genetic strings. The US National Marrow Donor Program (NMDP) has identified in excess of 74,000 different haplotypes, based on testing of its donor base when five loci (A, C, B, DRB1, and DQB1) are considered (Gragert et al., 2013; National Marrow Donor Program, 2015). Since a person's phenotype is based on two haplotypes, there are potentially $\binom{74,000}{2}$ or roughly 2.7B possible phenotypes, based strictly on estimates derived from alleles that have been observed in the donor population, and potentially a great deal more. The NDMP data set lists frequencies for these haplotypes within five broad racial groups and 21 racial sub-groups. Haplotype frequencies listed range from 5.9E−2 to 3.6E−7, suggesting phenotype frequencies that range from 3.5E−3 to less than 1.3E−13 (National Marrow Donor Program, 2015).

2.2. Registry size and matching probability

Finding a match for a patient with a particular phenotype i in a registry of size n, involves a series of Bernoulli trials, each with the probability of success equal to the frequency of the phenotype f_i within a donor population. Thus, the probability of finding at least one match for a patient in a registry of n donors is one minus the probability of finding no matches (Schmidt, Suater, Pingel, & Gerhard, 2014):

$$p(n) = 1 - (1 - f_i)^n$$

Equation (1). The probability of finding at least one match for a patient with phenotype i in a registry of size n.

The likelihood of successfully matching a patient having phenotype frequency f_i in a registry of size n is related to both the rareness of the patient phenotype and the size of the registry, where the larger the registry, the greater the chance of finding a donor with a particular phenotype, and thus the rarer the phenotype frequency that can be reliably found. See Fig. 1 for an illustration of this concept. However, because the probability of a match asymptotically approaches 1.0, the law of diminishing returns applies; in a registry of 10,000 persons, recruiting a further 100 people is much more likely to add an individual with an as-of-yet unseen genetic makeup, than adding 100 people to a registry of 1 M persons (Querol, Mufti, & Marsh, 2009). Accordingly, determining the optimal size of a stem cell registry is a question of some theoretical and practical interest.

In general, it is not possible to estimate phenotype frequencies directly from the population by sampling, due to the diversity of the human genotype. Accordingly, haplotype frequencies are estimated from phenotype information, rather than population samples. Hawley and Kidd (1995) describe a Fortran program to implement an instance of the expectation-maximization (EM) algorithm to estimate haplotypes and their frequencies, based on observed phenotype data. The algorithm starts with observed phenotype data and, using an iterative process, estimates the frequencies for haplotypes that would have maximized the likelihood of observing the data. Kollman, Abella, and Baitty (2005) similarly use an EM algorithm to determine haplotype frequencies based on data from the US NMDP registry to obtain ethnic specific predictions for finding a matched donor of a given size registry. An economic argument can be made that registry size is optimized when the marginal cost of adding an additional registrant becomes less than the expected value of the marginal gains to patients, but to do so requires an explicit cost be attached to not providing a match (Bergstrom et al., 2009). There is some evidence that an expansion to registries worldwide would be of potential value (Sheehan-Connor et al., 2015). However, economic evaluation is complicated by the fact that genetic diversity varies by ethnicity and thus the costs and benefits of recruiting an additional registrant are not uniform across all ethnic groups. Methods similar to that of Kollman et al., (2005) have been applied in a variety of jurisdictions for both adult registries (Buhler, Nunes, Nicoloso, Tiercy, & Sanchez-Mazas, 2012; Eberhard et al., 2010; Maiers, Halagan, & Joshi, 2014; Schmidt et al., 2009; Schmidt et al., 2011) and cord

Please cite this article as: J.T. Blake et al., Modeling the optimal ethnic composition of an adult stem cell registry, European Journal of Operational Research (2016), http://dx.doi.org/10.1016/j.ejor.2016.05.015