Published in final edited form as:
Comput Stat Data Anal. 2007 August ; 51(12): 5494–5506. doi:10.1016/j.csda.2007.02.030.

A matching algorithm for the distribution of human pancreatic islets

Dajun Qian[*], **John Kaddis**, and **Joyce C. Niland**
Division of Information Sciences and Administrative and Bioinformatics Coordinating Center for the Islet Cell Resource Center Consortium, City of Hope National Medical Center, Duarte, California, USA

Abstract

The success of human pancreatic islet transplantation in a subset of type 1 diabetic patients has led to an increased demand for this tissue in both clinical and basic research, yet the availability of such preparations is limited and the quality highly variable. Under the current process of islet distribution for basic science experimentation nationwide, specialized laboratories attempt to distribute islets to one or more scientists based on a list of known investigators. This Local Decision Making (LDM) process has been found to be ineffective and suboptimal. To alleviate these problems, a computerized Matching Algorithm for Islet Distribution (MAID) was developed to better match the functional, morphological, and quality characteristics of islet preparations to the criteria desired by basic research laboratories, i.e. requesters. The algorithm searches for an optimal combination of requesters using detailed screening, sorting, and search procedures. When applied to a data set of 68 human islet preparations distributed by the Islet Cell Resource (ICR) Center Consortium, MAID reduced the number of requesters that a) did not receive any islets, and b) received mis-matched shipments. These results suggest that MAID is an improved more efficient approach to the centralized distribution of human islets within a consortium setting.

Keywords

Islet distribution; matching algorithm; exhaustive search; space reduction; importance sampling

1. Introduction

The inability to respond to or produce insulin results in diabetes and leads to potentially severe secondary complications in individuals with this disease (DCCT, 1993; Nathan, 1993; Pinto et al., 2004; Taylor et al., 2004; Weir and Bonner-Weir, 2004). The malfunction and/or destruction of the pancreatic beta cell play a key role in the progression of diabetes. Beta cells are responsible for the production and storage of insulin, the hormone that controls blood glucose levels, and are one of four main cell types in the pancreas clustered together to form what are known as “islets of Langerhans” (Hardikar, 2004; Ahren and Taborsky, 2003).

---

[*] Corresponding author. Tel.: +1 626 2564673, Ext. 62685; fax: +1 626 4717106. dqian@coh.org (D. Qian), jkaddis@coh.org (J. Kaddis), jniland@coh.org (J.C. Niland).

Emerging strategies for the treatment of diabetes include islet replacement and/or renewal therapies (Ramiya, 2004; Trucco, 2005). Islet replacement strategies for a subset of type 1 diabetic patients are predominately based on the use of pancreata from a limited supply of human cadaveric donors (Nanji and Shapiro, 2006). The use of specialized pancreas processing facilities is required in order to generate high quality human islets derived from donated cadaveric pancreata. Organ procurement, islet isolation, final batch preparations, and product release characterization are highly complex procedures involving a chain of intricate processing steps. Relatively few centers around the world are capable of providing human islets for clinical transplantation purposes, yet the need for such islets is growing rapidly. It is estimated that between 1999 through the middle 2004, 43 institutions worldwide processed human pancreata for islet transplantation in over 470 recipients (Gaglia et al., 2005).

As a result of the improved outcomes demonstrated in patients receiving islet transplants (Shapiro et al., 2000), the need for human islets has increased in both clinical and basic research settings (Berney et al., 2005; Ridgway et al., 2005). With the hope that improvements in human islet isolation and transplantation will lead to better treatments and durable insulin independence for type 1 diabetes, the National Islet Cell Resource Center (ICR) consortium was established in 2001 (Knazek, 2002). The ICRs are a group of academic laboratories across the United States charged with providing human pancreatic islets for clinical and basic research purposes.

One of the key responsibilities of the ICR Administrative and Bioinformatics Coordinating Center (ABCC) is to coordinate the islet distribution activities of the ICR basic science Human Islet Distribution program. The existing islet distribution approach allows each producer, i.e. ICR laboratory, to make local decisions about who receives islets based on a centrally provided ABCC list of requesters, i.e. ICR approved investigators. Under this local decision making (LDM) approach, producers are often limited in their ability to closely match the characteristics of available islets to the desired criteria described in requesters' applications. Thus, requesters with identical needs may not receive a comparable number of islets, if any at all. The LDM approach, however, is used because it is often the quickest way to distribute islets in the least amount of time, ideally within 24 hours of production, thus minimizing the degradation of functional viability.

The problem of equitably allocating a limited supply of islets to multiple requesters in a limited amount of time can be viewed as a problem in combinatorial optimization (Nemhauser, 1988). The classic binary knapsack problem seeks to find an optimal set of objects that will fit into a knapsack of fixed size and volume, while simultaneously considering the unique attributes of each object (Martello and Toth, 1990). However, the problem of matching islet offers differs somewhat in that the parameters of a requesting investigator vary, rather than representing a fixed set of known attributes.

Moreover, this variation is one of several factors to consider when attempting to match islets generated by a producer to approved requesters. Not only do the characteristics of desired islets differ based on the intended research purpose of each requester, the shortage of organ donation and the under-utilization of pancreata also create an unpredictable supply of human islets that often falls short of demand at any point in time. Furthermore, the variable quality of available organs and the differing pancreas processing procedures used by a given producer of islets cause the resulting islets to vary in amount, purity and viability. To preserve maximum islet quality, islets must be shipped immediately following the completion of pancreatic processing, usually within 24 hours, to minimize degradation of the islets over time. Therefore the proximity of requesters to producers must be taken into account, along with considerations such as study start and end dates, and frequency, quality, and quantity of desired islets for experimentation.

A major challenge is that such an islet distribution algorithm should not only be computationally efficient, but also aid in the regulatory and functional oversight activities of the ICR consortium. These include the ability to decrease the priority of a requester who has previously rejected a valid islet offer, and increase the priority of certain investigators conducting research deemed critical by the ICR funding agencies. The preferences of a requester to receive islets from a selected producer can be based on the proximity, ongoing collaborative relationship, and/or the perceived quality of islets by the producer.. In addition to these fairly subjective criteria, the algorithm must take into account the objective criteria of islet purity (minimum and ideal), viability (minimum and ideal), quantity (minimum and ideal) and frequency (for the duration of the scientific project) of shipments for all approved studies.

To accommodate all of these objective and subjective criteria, a matching algorithm for islet distribution (MAID) was developed and tested as a tool to optimize the islet distribution activities in a practical setting. In this paper we describe the development and testing of this islet matching algorithm. We first apply MAID using actual data to model ICR Distribution Program behavior, including ICR production activity over a fixed time period, and shipments of islets to approved investigators during this timeframe. We then apply the algorithm to simulated data, varying certain key conditions anticipated to impact the ability to successfully match islet cell offers from producers to the pool of waiting islet requesters.

2. Methods

2.1. MAID development

We consider an islet distribution program consists of I producers and J requesters. Each islet isolation, labeled u, is generated from a donated cadaveric pancreas organ and characterized by various parameters, including a) producer i generating the islets, b) date t islets are isolated, c) number of available islets Q, measured in unit of islet equivalent (IEQ), for distribution, and d) islet purity P and viability V measured as percentages from 0% to 100%. We omit the index u in each of the above isolation parameters for simplicity. Each requester j must submit an application for islets prior to approval, and following information within the application is used to classify the requester: a) date of application approval $t_{j,\mathrm{apv}}$, b) supplier preference vector to all producers $\{y_j(i) \mid 1 \le i \le I\}$, c) minimum days between islet shipments x_j, d) minimum and ideal numbers of IEQs per shipment $q_{j,\mathrm{min}}$ and $q_{j,\mathrm{ideal}}$, e) minimum and ideal purity values $p_{j,\mathrm{min}}$ and $p_{j,\mathrm{ideal}}$, and f) minimum and ideal viability values $v_{j,\mathrm{min}}$ and $v_{j,\mathrm{ideal}}$.

The MAID algorithm uses these variables to generate an optimal offer list (OOL) of requesters via three computational components: a) a screening analysis to identify all qualified requesters, b) a sorting analysis to score and rank qualified requesters, and c) a search procedure for selection of OOL under required conditions. Requesters on the OOL are contacted to confirm acceptance of the islet offer. If a requester rejects the offered islets, those islets will be used to generate an alternative OOL. Fig. 1 provides an algorithm schema of MAID under a consortium setting, and the sections below describe each computational component in more details.

2.1.1. Screening analysis for qualified requesters

When a batch of islets, characterized by parameters u, i, t, Q, P and V, are generated and available for distribution, a screening analysis is first performed to identify a list of all qualified requesters. Specifically, a requester $j \in \{1, \dots, J\}$ is qualified if all of the following criteria are met:

1. For a requester j who has received at least 1 islet shipment, the time since last shipment should be at least the minimum days between shipments desired by the requester, i.e., $t - t_{j,\mathrm{LS}} \geq x_j$, where $t_{j,\mathrm{LS}}$ is the date of last shipment to requester j.
2. Producer i is an acceptable supplier to requester j, i.e., $y_j(i) = 1$.
3. The number of available IEQs is at least the minimum number of IEQs per shipment desired by requester j, i.e., $Q \geq q_{j,\min}$.
4. The purity and viability of available islets are not lower than the corresponding minimal acceptable values, respectively, i.e., $P \geq p_{j,\min}$ and $V \geq v_{j,\min}$.

For ease of notation, all qualified requesters $\{j_1, \dots, j_K\} \subseteq \{1, \dots, J\}$ identified in screening analysis will be indexed as $k = 1, \dots, K$ in subsequent analyses.

2.1.2. Sorting requesters by priority scores

We next compute a priority score s_k for each qualified requester k using a form of

$$s_k = w_k (a_k b_k c_k d_k) + e_k, \tag{1}$$

and sort the K requesters in descending order of their scores to generate an ordered list of candidates to be included in the OOL. In order to appropriately incorporate the priorities mandated by the ICR Steering Committee and funding agencies, the coefficient terms in equation (1) are determined as follows:

1. w_k = waiting days of requester k, in addition to the minimum days between shipments as stated in requester's application. Specifically, w_k equals to $t - t_{k,\mathrm{apv}}$ if no offer has ever been made to requester k, or $t - (t_{k,\mathrm{LS}} + x_k)$ if the last offer on date $t_{k,\mathrm{LS}}$ was accepted and shipped, or $t - t_{k,\mathrm{LO}}$ if the last offer on date $t_{k,\mathrm{LO}}$ was rejected.
2. $a_k = 1.1$ if the distance between producer and requester allows same day delivery, or 1 if next day delivery is required.
3. $b_k = 1.1$ if the requester has peer reviewed funding support, or 1 if no such funding.
4. $c_k = 1.5$ if the purity of available islets is within 5% difference from the ideal value desired by the requester, or 1 if otherwise.
5. $d_k = 1.5$ if the viability of available islets is within 5% difference from the ideal value desired by the requester, or 1 if otherwise.
6. $e_k = \max(w_1, \dots, w_K)(1.1\times1.1\times1.5\times1.5)$ if the requester has preferred priority status to receive islets, or 0 if the requester has standard priority status.

The rationale behind equation (1) reflects several considerations. First, waiting time is considered as a baseline value in quantifying the priority scores among all qualified requesters. The choice of $w_k = t - t_{k,\mathrm{LO}}$ indicates that the waiting time will be reset to 0 at the time of rejecting a matched offer. As a side note, the dates of last shipment and last offer are equal (i.e., $t_{k,\mathrm{LS}} = t_{k,\mathrm{LO}}$ for requester k) if the last offer has been accepted and shipped, or the date of last shipment is earlier (i.e., $t_{k,\mathrm{LS}} < t_{k,\mathrm{LO}}$) if the last offer has been rejected. Second, coefficients a_k and b_k for delivery distance and funding status are chosen to increase the scores by a small 10% under favorable conditions. These two coefficients are found to have little or no influence on the resulting proportion of unmatched islets. Third, coefficients

$c_k = 1.5$ and $d_k = 1.5$ for purity and viability matches are chosen so as to minimize the proportion of unmatched islets. And fourth, the preferred priority status is considered as a dominant factor that allows a small portion of requesters to have the highest priority scores when so designated by the funding agencies.

2.1.3. Search for optimal offer list (OOL)

The search for OOL is to evaluate all possible subsets among the K qualified requesters obtained in screening analysis, and to identify the optimal one under given conditions. The screening analysis generally yields multiple potential requesters, and the final OOL is required to consist of up to 10 requesters who will be offered IEQs from a given isolation in order to utilize all the islets. We search the OOL under following three conditions: A) the number of requesters in optimal solution is less than or equal to $N_{\max}$, where $N_{\max} = 10$ is the default value and $1 \le N_{\max} \le 9$ are the alternative choices when needed, B) the number of unmatched islets is minimized, and C) the average of priority scores is maximized. When K is small, we identify the optimal solution based on an exhaustive search. Alternatively, when K is large, we search the optimal solution using a semi-exhaustive procedure, coupled with importance sampling and extended local search for improved search performance.

Several notations are needed before describing the proposed search procedure. Let $\Omega = \{1, \ldots, K\}$ denote the full list of K qualified requesters sorted in descending order of their priority scores, $\Omega_N = \{1, \ldots, N \mid N \le K\}$ denote a subset of Ω with the N top-score requesters in descending order of scores, and $\Omega_{(N)} = \{k_1, \ldots, k_N \mid N \le K\}$ denote a subset of Ω with any N requesters in descending order of scores. Of note, the number of matched IEQs in any offer list $\Omega_{(N)}$ must equal to or less than the number of available IEQs Q.

We begin by describing the search procedure under a simple scenario when the number of qualified requesters K is small. Let $M = \max(10, N_{\max})$ denote a small number for which exhaustive search is feasible. When $K \le M$, an exhaustive search in the full space of $2^K - 1$ subsets is performed to retain any subsets $\Omega_{(N)}$ such that $N \le N_{\max}$ and $\Omega_{(N)}$ results in 0 or minimal unmatched islets. The OOL is chosen to be the retained $\Omega_{(N)}$ with maximal average score, or a random one when multiple retained $\Omega_{(N)}$ have the same maximal average score.

We now describe the search procedure under an alternative and general scenario when the number of qualified requesters is large (i.e., $K > M$), and an exhaustive search among an exponential number of subsets is computational infeasible. A semi-exhaustive search procedure, coupled with importance sampling and extended local search, is used to identify the OOL under the required conditions A-C. The search procedure can be described in 3 steps as follows.

Step 1: Reduce Search Space to a Manageable Size—We obtain a reduced search space Θ_M containing M requesters for the search of OOL in two mutually exclusive situations. Let $Q(\Omega_N) = \Sigma_{k\varepsilon\Omega_N} q_{k,\text{ideal}}$ and $Q(\Omega_{(N)}) = \Sigma_{k\varepsilon\Omega_{(N)}} q_{k,\text{ideal}}$ denote the sums of ideal IEQs per shipment desired by requesters in Ω_N and $\Omega_{(N)}$, respectively. In a common situation that $K > M$ and $Q(\Omega_M) \ge Q$, the reduced search space is chosen to be all subsets in M top-score requesters, i.e., $\Theta_M = \{\Omega_{(N)} \mid \Omega_{(N)} \subseteq \Omega_M, N \le M\}$, and go to *Step 2*. In practice, the condition $Q(\Omega_M) \ge Q$ is expected to hold in more than 90% occasions because the number of IEQs Q generated from a cadaveric pancreas organ is rarely sufficient for offering 10 or more requesters.

In a less common situation that $K > M$ and $Q(\Omega_M) < Q$ (e.g., the M top-score requesters all desire very small numbers of IEQs per shipment), an importance sampling routine is employed to find a reduced space $\Theta_M = \{\Omega_{(N)} | \Omega_{(N)} \subseteq \Omega^*_{(M)}, N \le M\}$, where $\Omega^*_{(M)}$ is a set of M

selected requesters with the sum of their ideal IEQs per shipment equal to or greater than the number of available IEQs (i.e., $Q(\Omega^*_{(M)}) \geq Q$. To do this, we choose an initial $\Omega_{(M)}$ to be the M top-score requesters $\Omega_M = \{1, \ldots, M\}$, and then update $\Omega_{(M)}$ iteratively by substituting one requester $k \; \varepsilon \; \Omega_{(M)}$ with another one $l \; \varepsilon \; (\Omega - \Omega_{(M)})$ based on assigned sampling probabilities $\{f_k, k \; \varepsilon \; \Omega_{(M)}\}$ and $\{g_l, l \; \varepsilon \; (\Omega - \Omega_{(M)})\}$, respectively. Numerically, we choose f_k in proportion to $2^{-(M-k+1)}$ for each $k \; \varepsilon \; \Omega_{(M)}$ and g_l in proportion to $s_l \times q_{l,\text{ideal}}$ for each $l \; \varepsilon \; (\Omega - \Omega_{(M)})$, where k in quantity $2^{-(M-k+1)}$ represents requester k has the k-th largest score in $\Omega_{(M)}$. Of note, these sampling probabilities are chosen to force $\Omega_{(M)}$ in favoring requesters with both higher scores and larger requested islet amounts per shipment. If $Q(\Omega_{(M)}) \geq Q$ is reached within 1000 iterations, we choose $\Omega^*_{(M)} = \Omega_{(M)}$ to form a reduced search space Θ_M, and go to *Step 2*. Otherwise, we choose $\Omega_{(M)}$ at iteration 1000 as a temporary offer list, and go to *Step 3*.

Step 2: Exhaustive Search in Reduced Search Space Θ_M—If a reduced search space Θ_M is found in *Step 1*, we perform exhaustive search by evaluating a maximum of $2^M - 1$ subsets in Θ_M and choose an optimal one Ω^* under the required conditions A-C. If the identified Ω^* results in zero unmatched islets, we choose the OOL equals to Ω^*, and stop the search procedure. Otherwise, if Ω^* results in non-zero unmatched islets, we choose Ω^* as a temporary offer list, and go to *Step 3*.

Step 3: Extended Local Search When Needed—Occasionally, we may either fail in *Step 1* in finding a reduced search space, or fail in *Step 2* in identifying an optimal subset with zero unmatched islets. In such cases, an extended local search in the full space of all K qualified requesters is performed to identify an OOL under the required conditions A-C. To do this, we start with a temporary offer list $\Omega_{(N)}$ containing $N \leq M$ requesters as obtained in *Step 1*, or *Step 2*, when appropriate. We then substitute 1 requester a time in $\Omega_{(N)}$ with possible combinations of 1 to $M - N + 1$ requesters in $(\Omega - \Omega_{(N)})$, retain all subsets with minimal unmatched islets, and update the offer list $\Omega_{(N)}$ when substitutions are done for all requesters in $\Omega_{(N)}$ and conditional on the minimal unmatched islets. If at least one subset has zero unmatched islets, we choose the OOL to be the one with maximum average score, and stop the search procedure. Otherwise, if no subset has zero unmatched islets, we choose the OOL to be the one with minimum unmatched islets, and stop the search procedure. Of note, one loop by substituting each requester once in a corresponding offer list $\Omega_{(N)}$ is generally sufficient in above local search analysis, because experimental evaluations indicated no change for the resulting OOL in additional loops.

Infrequently, the available islets are not acceptable to any requesters (i.e., $K = 0$), or the quantity of available islets is more than the sum of ideal amounts per shipment desired by all qualified requesters or the $N_{\max}$ ones requesting the largest amounts per shipment. In such cases, all or a portion of islets can not be matched, and those unmatched islets are offered to any approved requesters willing to accept them with no influence on future offers.

2.2. Actual and simulated data sets

2.2.1. Actual ICR consortium data—We evaluated the proposed matching algorithm using a retrospective pilot data set of 68 human pancreatic islet isolations distributed under a LDM model and collected by the ICR consortium during January – September 2005. Within the 68 islet isolations, a total of 6,653,944 IEQs were shipped in 217 shipments from 8 producers to 62 requesters located in 18 states across the United States (Fig. 2). As can be seen in Fig. 3, certain requesters received more than the desired quantity of islets, while others received no islets over a given time period, even though their requesting criteria and other conditions may be similar. Among the 62 requesters, 23 (37%) were on requesters list for the entire 9 months, and the remaining 7 (11%), 26 (42%) and 6 (10%) requesters were added into the list during the first, middle and last 3 months, respectively. The islet supply versus demand ratio, defined as the ratio of total produced islets by all producers to total minimum requested islets by all requesters, in the 9 months period was 0.68.

To apply the MAID algorithm to these 68 isolations, we estimated rejection rate and imputed missing values as follows. The rate at which requesters rejected islet offers was estimated as 30%, based on an ongoing ICR consortium study during January – June 2006 (details not shown), and an hypothetical rejection rate of 15% was also tested to determine the impact of this factor. Missing values in 17 (27%) requesters each with at least one missing data point were imputed as follows to ensure complete data on all requesting criteria: i) missing values for desired minimum days between shipments were imputed by the mode value of 30 days, ii) no missing value existed for supplier preference data, iii) for paired minimum and ideal numbers of IEQs per shipment, mode values of 5,000 and 10,000 IEQs were used, respectively, if both values were missing, and ± 50% of the stated value was used if one value was missing, iv) for paired minimum and ideal purity values, mode values 0.50 and 0.90 were used, respectively, if both values were missing, and ± 50% of the stated value was used if one value was missing, and v) the imputation for paired minimum and ideal viability values was done the same way as the paired purity values. We also assume islets from rejected offers can be redistributed to other qualified requesters by running the algorithm repeatedly for a maximum of 5 times, and the unmatched islets also include rejected islets from the last algorithm run.

2.2.2. Data simulation—Data simulation mimicked a consortium for islet distribution consisting of 8 producers and 80 requesters in a 1 year period. Each simulation corresponds to 2 source data sets, one for 80 requesters and the other for a varied number of isolations during the study period. Both requester and isolation variables were simulated based on the corresponding data distributions in above mentioned consortium pilot data set. Specifically, the variables for each requester *j* were simulated as follows:

1. Date of application approval $t_{j,\mathrm{apv}}$ follows a density distribution so that the length of time in study is 1 year in 40 requesters and a random number of 0.5 to 1 year in remaining 40 requesters.
2. Minimum days between shipments x_j follows a truncated log-normal distribution with median (range) of 21 (7, 243) days.
3. Supplier preference vector $\{y_j(i) \mid 1 \le i \le I\}$ follows a density distribution so that the number of acceptable producers is 1, 2–5 and 8 in 16%, 29% and 55% requesters, respectively.
4. Ideal number of IEQs per shipment $q_{j,\mathrm{ideal}}$ follows a truncated log-normal distribution with median (range) of $2{\times}10^4$ (10^3, $5{\times}10^5$) IEQs. The corresponding minimum number of IEQs per shipment $q_{j,\min}$ equals to $q_{j,\mathrm{ideal}}$, $0.75{\times}q_{j,\mathrm{ideal}}$ and $0.5{\times}q_{j,\mathrm{ideal}}$ in 25%, 35% and 40% requesters, respectively.
5. Ideal purity $p_{j,\mathrm{ideal}}$ follows a truncated normal distribution with median (range) of 0.85 (0.50, 0.90). The corresponding minimum purity $p_{j,\min}$ equals to $p_{j,\mathrm{ideal}}$, $0.9{\times}p_{j,\mathrm{ideal}}$ and $0.8{\times}p_{j,\mathrm{ideal}}$ in 33%, 33% and 34% requesters, respectively.
6. Ideal viability $v_{j,\mathrm{ideal}}$ follows a truncated normal distribution with median (range) of 0.90 (0.50, 0.99). The corresponding minimum viability $v_{j,\min}$ equals to $v_{j,\mathrm{ideal}}$, $0.9{\times}v_{j,\mathrm{ideal}}$ and $0.8{\times}v_{j,\mathrm{ideal}}$ in 33%, 33% and 34% requesters, respectively.
7. Priority status: 20% requesters have preferred status and the remaining 80% have standard status.

8. Delivery distance: 37% requesters are within same day delivery distance from 1 producer. All the remaining deliveries require next day shipment.
9. Funding support: 69% requesters have peer-reviewed funding status.

The variables for each islet isolation u were simulated as follows:

1. Producer i is a random number selected from 1 to 8.
2. Isolation date t follows a Poisson distribution conditional on an average isolation frequency determined by a prefixed islet supply versus demand ratio.
3. Number of produced IEQs Q follows a truncated log-normal distribution with median (range) of 7.7×10^4 (8×10^3, 10^6) IEQs.
4. Purity P and viability V follow truncated normal distributions with median (range) of 0.90 (0.50, 0.95) and 0.92 (0.70, 0.99), respectively.

We performed 10 replicated 1-year simulations under each of the three isolation frequency scenarios corresponding to the low, moderate and high islet supply versus demand ratios of 0.3, 0.6 and 0.9, respectively. The 10 replicated simulations represented a total of 7.0×10^7 IEQs in 657 isolations, or 1.4×10^8 IEQs in 1494 isolations, or 2.0×10^8 IEQs in 2001 isolations when the supply versus demand ratio was 0.3, 0.6, or 0.9, respectively.

3. Results

3.1. Application of MAID to actual ICR consortium data

By comparison to the LDM approach in place during the time period tested, the MAID-derived results revealed several advantages over those in actual distribution data (Table 1 and Fig. 3). First, in actual distributions, 24% islets (1,576,498 IEQs) in 35 shipments met the 4 matching criteria on producer preference, minimum IEQs per shipment, minimum purity and minimum viability, 24% islets (1,578,656 IEQs) in 70 shipments did not meet at least 1 criterion, and 53% islets (3,498,790 IEQs) in 112 shipments contained missing data on islet characteristics and/or receivers' criteria. When the proposed MAID algorithm was applied to this data set after missing data were imputed as described above and under an estimated 30% rejection rate, 91% islets (6,038,048 IEQs) in 194 shipments met all the 4 criteria, and the remaining 9% (615,896 IEQs) were either unmatched or rejected islets. As separate analyses, the total matched islets increased to 97% under both rejection rates of 15% and 0%, respectively. Second, of the 62 requesters in study, 16 (26%) did not receive any islets in actual distribution activities and only 3 (5%) did not received any matched islets in MAID-derived results ($p = 0.001$ by Fisher's exact test, Fig. 3). For the 56 requesters on study for at least 3 months, the above numbers became 11 (20%) and 1 (2%) in actual and MAID-derived results, respectively ($p = 0.004$). Lastly, the number of requesters who received more islets than desired ideal amounts was 13 (21%) in actual activities and 0 (0%) in MAID-derived results ($p = 0.001$, Fig. 3).

3.2. Application of MAID to simulated data

Several statistics in the combination of 10 replicated 1-year simulations were displayed under islet supply versus demand ratios of 0.3, 0.6 and 0.9, respectively, to show both the patterns of relationship and the variations in simulated data sets (Fig. 4). Fig. 4a indicates the number of qualified requesters in all algorithm runs for all isolations had a negative correlation with the magnitude of islet supply versus demand ratio (Pearson correlation $r = -0.30 \pm 0.08$ (range -0.14 to -0.43), $p < 0.0001$ in each of the 10 simulations). In particular, 5 or more qualified requesters were found in 82%, 69% and 62% isolations when islet supply versus demand ratios were 0.3, 0.6 and 0.9, respectively. For islet supply versus demand ratio of 0.6 or less, the number of qualified requesters was large enough in more than 67% algorithm runs that allowed successful search of OOL with no leftovers, and was zero in less than 1% runs that stopped the algorithm for further searches.

Fig. 4b indicates the proportion of unmatched islets was increased from 5% to 11% to 18% as a result of an increase in islet supply versus demand ratio from 0.3 to 0.6 to 0.9, respectively. The proportion of unmatched islets had a 16% to 32% difference when comparing isolations with both purity and viability of 0.75 or higher to the remaining ones with low purity and/or viability. In particular, the proportion of unmatched islets was less than 7% in isolations with both purity and viability of 0.75 or higher when the supply versus demand ratio was 0.6 or less.

Fig. 4c indicates requesters with preferred priority status received higher portions of requested islets than others with standard priority status. When islet supply versus demand ratio was at a moderate level of 0.6, the proportions of received islets among requested ideal quantity in requesters with and without preferred priority status were 0.56 ± 0.24 and 0.43 ± 0.24, respectively. Additionally, requesters accepting lower quality islets (e.g., $p_{j,\min} < 0.75$ and/or $v_{j,\min} < 0.75$) received a 10% to 16% higher portion of requested ideal quantity than others only accepting high quality islets. Similarly, requesters asking for less than 10^6 IEQs per year received a 13% to 23% higher portion of requested ideal quantity than others requesting 10^6 or more IEQs per year.

For islet supply versus demand ratio of 0.6, higher rejection rates resulted in higher proportions of unmatched islets (Table 2 and Fig. 5b). Numerically, the proportions of unmatched islets were 4.5% (0.68 ± 0.53 10^6 IEQs per year), 7.8% (1.16 ± 0.53 10^6 IEQs per year) and 11.4% (1.67 ± 0.53 10^6 IEQs per year) under rejection rates of 0, 0.15 and 0.30, respectively ($p < 0.0001$ for difference in 10 simulations by one-way analysis of variance [ANOVA], Table 2). The number of qualified requesters in the initial algorithm runs in all isolations was 17.0 ± 10.0 under rejection rate of 0, and dropped to 14.4 ± 9.1 and 13.8 ± 9.0 under rejection rates of 0.15 and 0.30, respectively ($p = 0.01 \pm 0.01$ with range from <0.0001 to 0.04 in 10 simulations, Fig. 5a). The number of qualified requesters in second and later algorithm runs further dropped to 7.9 ± 6.9 and 7.7 ± 7.0 under rejections of 0.15 and 0.30, respectively. Interestingly, the average proportion of received IEQs among requested ideal quantity in all requesters was 0.41 ± 0.23 under rejection rate of 0, and slightly higher at 0.45 ± 0.24 and 0.46 ± 0.24 under rejection rates of 0.15 and 0.30, respectively ($p = 0.43 \pm 0.21$ with range from 0.10 to 0.69 in 10 simulations, Fig. 5c). Requesters asking for small number of IEQs per shipment tended to have more chance to receive offers from the redistribution of rejected islets in second and later algorithm runs.

4. Discussion

In 1998, the Juvenile Diabetes Research Foundation International (JDRFI) created the JDRF human islet distribution program. As the first of its kind at the time, the program sponsored nine pancreas processing facilities worldwide in the United States, Europe, and Canada to distribute human pancreatic islets to approved investigators conducting research in the a) prevention of type 1 diabetes, b) restoration and maintenance of normal blood glucose, and/or c) prevention and treatment of complications. Each facility was responsible for independently managing a centrally provided JDRF list of approved investigators and ensuring that requesters on that list received islets in a timely fashion. Upon discontinuation of that program, the ICR distribution program was created shortly thereafter to address the need for human islets in the diabetes research community. This program adapted the JDRF's LDM approach of allowing each pancreas processing facility to determine who receives islets, using an ABCC provided list of investigators conducting ICR approved research.

The LDM approach to islet distribution has advantages and disadvantages. The immediate benefit to this approach is that producers can quickly make their own decisions about who receives the islets. These decisions will also include determining a) the amount and type of islets each requester receives, b) reasons for excluding or including requesters, and c) how the islets are packaged and shipped. The LDM approach also has serious limitations. First, because each producer does not know the islet offer and shipment history of requesters by other producers, certain requesters may receive more islet offers than others, while some may be excluded altogether from being contacted. Next, preferential treatment of certain requesters by a producer is always a possibility, and especially a concern to those requesters located outside of the producers institution. Finally, requesters that need sequential islet shipments in a compressed period of time to perform a series of experiments must coordinate each request with multiple producers to fill the demand. The development of MAID was needed to address the shortcomings of the LDM approach, and to obtain an optimal solution for matching producer islets to requester needs by using all available information.

MAID is designed to optimize the islet distribution process in a setting of multiple producers with multiple approved requesters eagerly awaiting islets. The algorithm searches for an OOL of requesters for a given amount of islets using a screening analysis to identify all qualified requesters, a sorting analysis to rank qualified requesters in descending order of priority scores, and a semi-exhaustive search procedure for the selection of optimal requesters that minimizes the amount of unmatched islets and maximizes the average of priority scores. A screening analysis may also include other criteria not evaluated in this report, such as those based on matching for type of islets (e.g., fresh, cultured, cryopresserved) and other penalties for rejections of previous offers. In particular, if the number of requesters on a waiting list is constantly large (e.g., 30 or more) and some requesters have not received islet offers for excessively long periods, we may impose a suspension period for rejecting matched offers in order to control the size of waiting list.

The considerations for implementing the semi-exhaustive search procedure are two fold. First, exhaustive search in the full space of $2^K - 1$ subsets is computational infeasible when the number of qualified requesters K is large. Second, the OOL obtained from the semi-exhaustive approach is guaranteed to represent the global optimum if the following two common conditions are met: a) an exhaustive search in reduced space is performed without resorting to the subroutines of importance sampling and extended local search, and b) the resulting OOL corresponds to zero unmatched islets. Therefore, an estimate of more than 80% OOLs identified in various matching analyses reported in sections 4 and 5 were representing the global optimal solutions. Woodruff and Reiners (2004) have discussed the utilization of heuristic optimization to find good solutions in a reasonable amount of time, applying data mining for algorithm development. We have taken a similar approach here to more optimally allocate the limited supply of islets to multiple requesters, using a semi-exhaustive search procedure for combinatorial optimization.

Other forms of optimization algorithms could be considered for approaching this real world problem of matching islet cell isolations to requesters, while incorporating both the required subjective and objective criteria. For example, a tree pruning (TP) algorithm could be used to search for an optimal combination of requesters by removing non-optimal branches in a systematic fashion. The TP technique is efficient, well developed, and expected to have an equal or better performance than our brute force “local search” approach. However, we believe that the possibility of obtaining an improved OOL solution using a TP search is very low for the following reasons. First, because more than 50% of the requesters are willing to have a flexible range for amount of IEQs per shipment (i.e., $q_{j,\min}$ typically is 20% to 50% less than their $q_{j,\text{ideal}}$), the chance of finding a global optimal solution among the top 10 scoring requesters is high (normally >90%). Therefore, the extended searches via TP or others generally would be unnecessary. Second, on the same basis (i.e., $q_{j,\text{min}} < q_{j,\text{ideal}}$), our local search approach is easy to implement, and is expected, although not guaranteed, to rapidly reach the optimal solution in most cases. If the amounts of IEQs per shipment were unchangeable (i.e., $q_{j,\text{min}} = q_{j,\text{ideal}}$) for all requesters, then a TP search would represent a clear improvement over our proposed algorithm; however the unmatched islets (i.e., leftovers) would be expected to occur much more often for all search procedures. Third, our algorithm does not allow a combination with "N over 10 qualified matches". For islets produced in one isolation, we want to find a combination with N not to exceed 10 offers. Therefore the benefit of applying a TP routine would be trivial, unless a) the criterion of $q_{j,\text{min}} = q_{j,\text{ideal}}$ is true in most requesters, and/or b) N = 20 or more offers are allowed for some "large" isolations, neither of which hold true in our experience.

We have demonstrated by application to real/imputed and simulated data that our islet matching algorithm improves the ability of the ICR consortium to rapidly and efficiently match available islets to the list of approved requesters. When applying MAID to a real data set consisting of 68 isolations, fewer unmatched islets remained than when using the LDM approach. We note a comparison of actual and MAID-derived results in this data set may be biased due to the existence of missing data, but the observed statistics clearly favor the MAID approach over the LDM model. Simulation studies also confirmed the benefits of the MAID approach, as reflected in both the proportion of unmatched islets in all isolations and the proportion of received among requested islets in all requesters. Additionally, MAID can incorporate new rules in each computational component, including the screening analysis, priority score calculations, and search for OOL. Islets offer rejections can be redistributed to other qualified requesters by running MAID repeatedly. If, however, the amount of available islets at the time of shipment is less than the initial amount used in obtaining the OOL, some requesters in the original offer list will not receive islets. Such exclusions changes should be reflected in future islet distribution analyses.

Operation of the matching algorithm presents remaining challenges, including improving the supply versus demand ratio and the uniform quality of islets produced. Furthermore, as it was seen that half of all rejections are due to lack of contact, invoking a reliable rapid mode of response from requesters when an islet isolation offering is available is essential. A web-based application is currently underway to further evaluate the feasibility and efficiency of this approach.

In conclusion, our results indicate that a centralized consortium of multiple producers is able to match more islets to more requesters than the situation when islets are distributed separately, and the proposed matching algorithm is a useful tool for islet distribution activities under a consortium setting.

Acknowledgments

Islet distribution data was obtained through the Islet Cell Resource Center (ICR) Basic Science Human Islet Distribution Program. Helpful discussions about the rules for the algorithm were provided by ICR Steering Committee members. The authors are also grateful to Nora Ruel for help with preparation of the manuscript and Doug Stahl and Rebecca Nelson for early discussions on the algorithm. This work was supported by a grant from the National Center for Research Resources (1U42-RR17673-01), a component of the US National Institutes of Health, Department of Health and Human Services.

References

Ahren, B.; Taborsky, GJ. Beta-Cell Function and Insulin Secretion. 6. McGraw-Hill; New York: 2003.

Berney T, Buhler LH, Morel P. Pancreas allocation in the era of islet transplantation. Transplant Intern. 2005; 18:763–767.

The Diabetes Control, Complications Trial Research Group. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. N Engl J Med. 1993; 329:977–986. [PubMed: 8366922]

Gaglia JL, Shapiro AM, Weir GC. Islet transplantation: progress and challenge. Arch Med Res. 2005; 36:273–280. [PubMed: 15925017]

Hardikar AA. Generating new pancreas from old. Trends Endocrinol Metab. 2004; 15:198–203. [PubMed: 15223048]

Knazek RA. The human pancreatic islet cell resource consortium. Diabetes Technol Ther. 2002; 4:551–552. [PubMed: 12396750]

Martello, S.; Toth, P. Knapsack problems: algorithms and computer implementations. 1. John Wiley & Sons, Inc; New York, NY USA: 1990.

Nanji SA, Shapiro AM. Advances in pancreatic islet transplantation in humans. Diabetes Obes Metab. 2006; 8:15–25. [PubMed: 16367878]

Nathan DM. Long-term complications of diabetes mellitus. N Engl J Med. 1993; 328:1676–1685. [PubMed: 8487827]

Nemhauser, GL.; Wolsey, LA. Integer and combinatorial optimization. 1. Wiley-Interscience; New York, NY, USA: 1988.

Pinto A, Tuttolomondo A, Di Raimondo D, Fernandez P, Licata G. Cerebrovascular risk factors and clinical classification of strokes. Semin Vasc Med. 2004; 4:287–303. [PubMed: 15630637]

Ramiya V, Schatz D. Islet replacement vs. regeneration: hope or hype? Pediatr Diabetes. 2004; 5(Suppl 2):45–56. [PubMed: 15601374]

Ridgway DM, White SA, Kimber RM, Nicholson ML. Current practices of donor pancreas allocation in the UK: future implications for pancreas and islet transplantation. Transplant Intern. 2005; 18:828–834.

Shapiro AM, Lakey JR, Ryan EA, Korbutt GS, Toth E, Warnock GL, Kneteman NM, Rajotte RV. Islet transplantation in seven patients with type 1 diabetes mellitus using a glucocorticoid-free immunosuppressive regimen. N Engl J Med. 2000; 343:230–238. [PubMed: 10911004]

Taylor GW, Manz MC, Borgnakke WS. Diabetes, periodontal diseases, dental caries, and tooth loss: a review of the literature. Compend Contin Educ Dent. 2004; 25:179–192. [PubMed: 15641324]

Trucco M. Regeneration of the pancreatic beta cell. J Clin Invest. 2005; 115:5–12. [PubMed: 15630433]

Weir GC, Bonner-Weir S. Five stages of evolving beta-cell dysfunction during progression to diabetes. Diabetes. 2004; 53(Suppl 3):S16–S21. [PubMed: 15561905]

Woodruff DL, Reiners T. Experiments with, and on, algorithms for maximum likelihood clustering. Comput Statist Data Anal. 2004; 47:237–253.

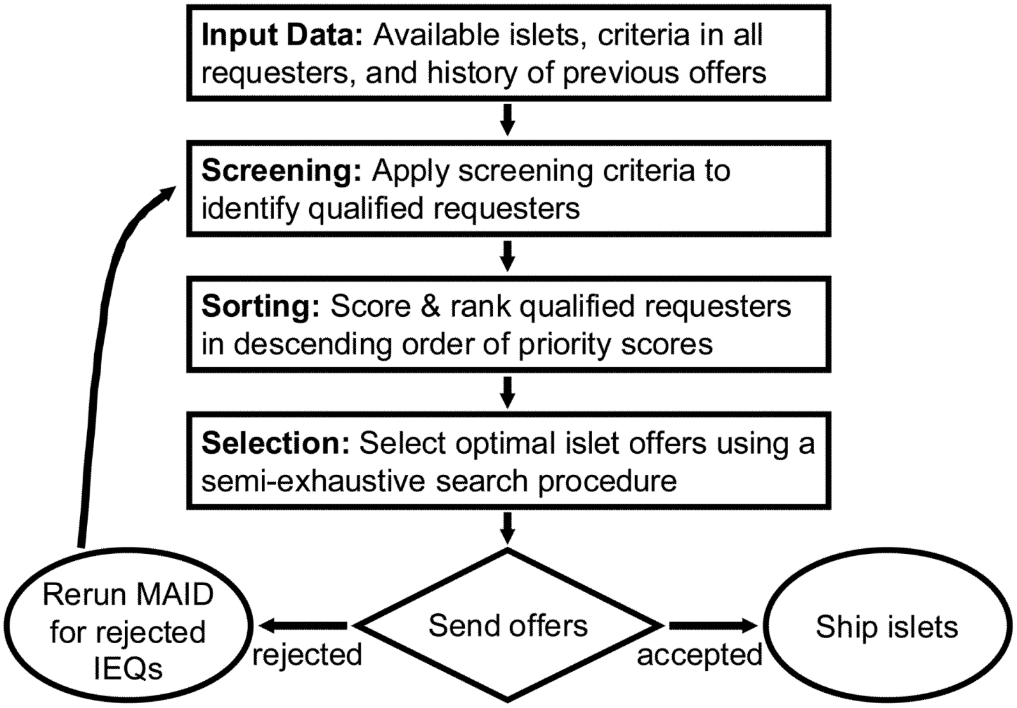

Fig. 1.

Algorithm schema of MAID.

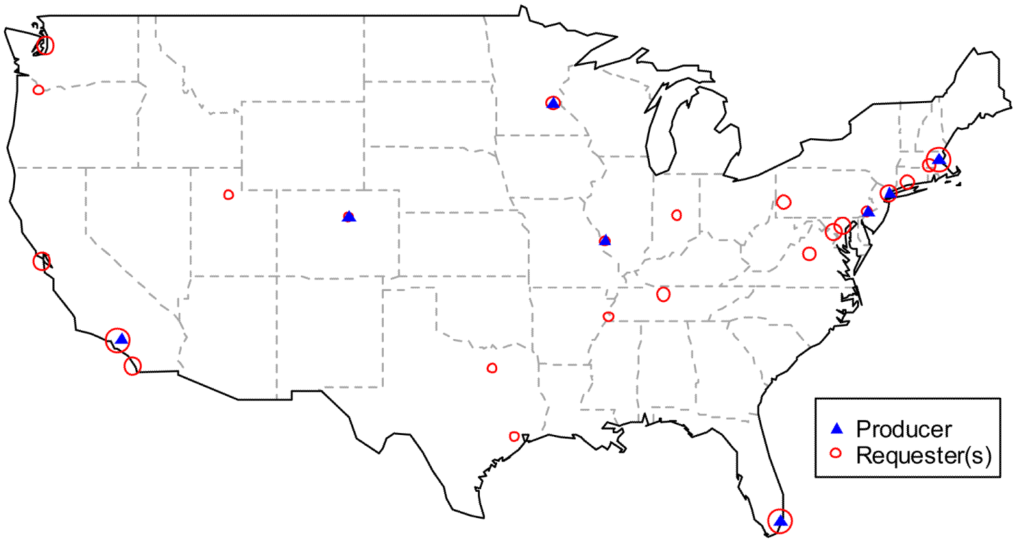

Fig. 2.

Geographical locations of 8 producers and 62 requesters in an ICR consortium study during January–September 2005. Each triangle represents an ICR producer, and each circle represents one or multiple requesters with its area proportional to the number of requesters at a same or very closed location.

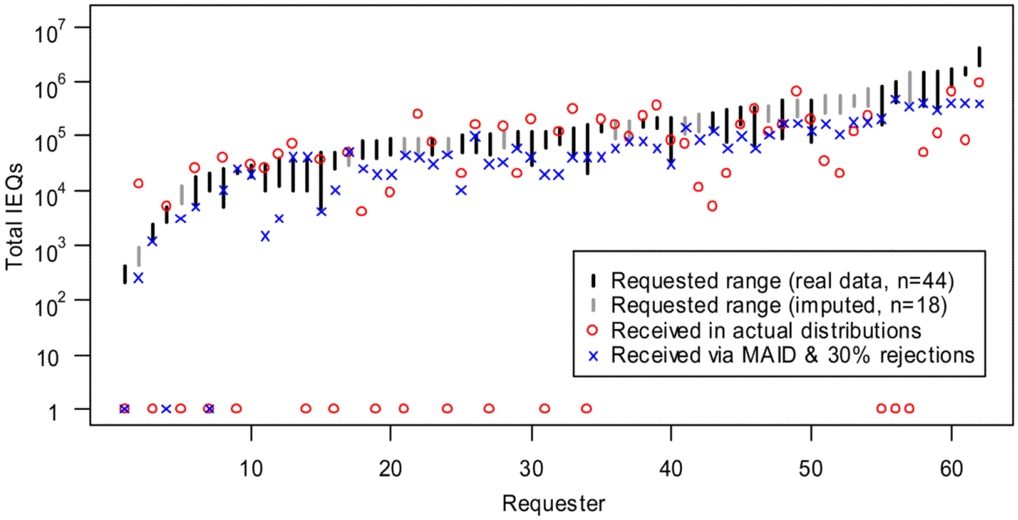

Fig. 3.

Islet distribution of 68 isolations for the ICR consortium data during January – September 2005 using the LDM model and after applying MAID. Each vertical solid line represents the range of total islet amount desired by a requester during the study period, which was estimated by the corresponding minimum and ideal amounts per shipment, minimum days between shipments and length of time in study. Each vertical gray line has the same meaning as a solid line, except that the range was imputed in combination of known data and population mode values due to missing values. Each cycle "○" represents the total islets received by a requester in actual distribution activities, and each cross "×" represents the total matched islets received by a requester in MAID-derived results under a 30% rejection rate in all offers.

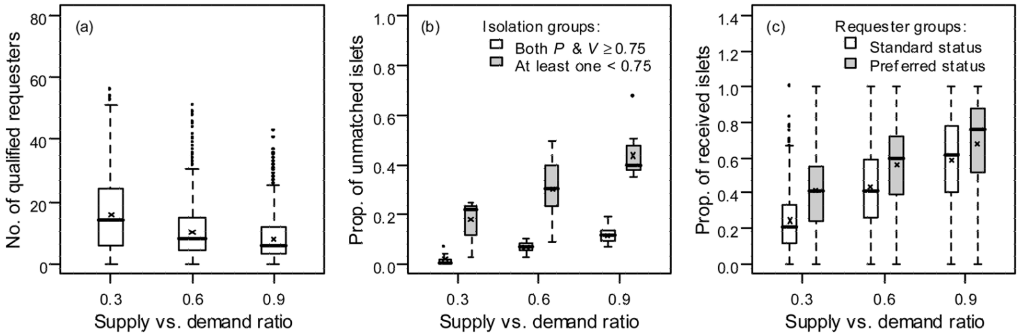

Fig. 4.

Simulation results by islet supply versus demand ratio. Each box-plot in each sub-figure represents a relationship in 10 replicated simulations. Each simulation corresponded to a consortium setting of 8 producers, 80 requesters, and a varied number of isolations distributed in a 1 year period and under a rejection rate of 30%. The 10 replicated simulations represented a total of 657, 1494 and 2001 islet isolations generated under islet supply versus demand ratios of 0.3, 0.6 and 0.9, respectively. (a) Number of qualified requesters in each algorithm run. (b) Proportion of unmatched islets stratified by isolation quality groups on purity and viability values. (c) Proportion of received islets among requested ideal quantity stratified by requester groups on priority status.

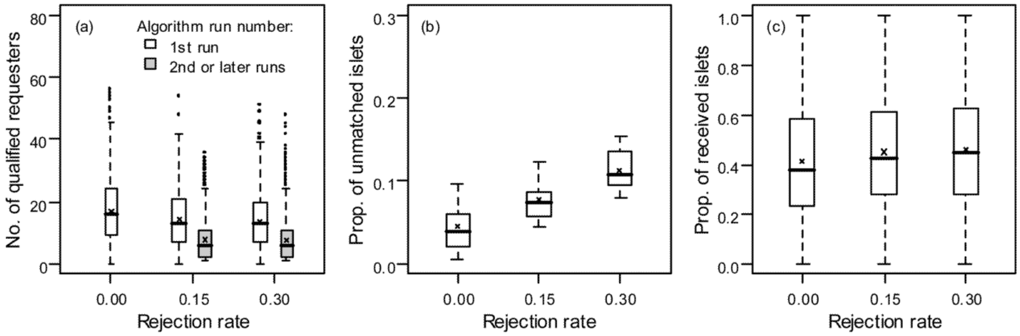

Fig. 5.

Simulation results by rejection rate. Each box-plot represents a relationship in 10 replicated simulations generated under islet supply versus demand ratio of 0.6, and the y-axis titles have the same meanings as in Fig. 4.

Table 1

Actual and MAID-derived results for the ICR consortium data during January–September, 2005

Variable	**Actual data**	**MAID under 15% rejections**	**MAID under 30% rejections**
For 62 requesters:			
No. of shipments received	2 (0, 17)	2 (1, 5)	3 (0, 8)
Proportion of requested shipments received			
0%	16 (26%)	0 (0%)	3 (5%)
1–49%	29 (47%)	58 (93%)	38 (61%)
50–99%	10 (16%)	3 (5%)	17 (27%)
100%	1 (2%)	1 (2%)	4 (6%)
>100%	5 (8%)	0 (0%)	0 (0%)
Unable to determine	1 (2%)	---	---
For 68 isolations:			
No. of algorithm runs per isolation	---	1.8 ± 1.2	2.3 ± 1.5
No. of qualified requesters per algorithm run	---	10.1 ± 10.2	10.6 ± 9.5
Total shipments for all isolations	217	202	194
Criteria not fully specified	3,498,790 (53%)	---	---
Criteria fully specified			
Matched	1,576,498 (24%)	6,438,944 (97%)	6,038,048 (91%)
Unmatched	1,578,656 (24%)	215,000 (3%)	615,896 (9%)

Descriptive statistics are median (range), mean ± standard deviation, or number (%), when appropriate. “Criteria not fully specified” corresponds to the shipments with missing data for evaluation of all 4 matching criteria on supplier preference, minimum IEQs per shipment, minimum purity and minimum viability. “Matched” represents the shipped islets met all above 4 matching criteria. “Unmatched” denotes the shipped islets did not meet at least 1 of the 4 criteria in actual data, or the combination of unmatched and rejected islets in the last algorithm runs in MAID-derived results. “---” = not applicable.

Table 2

MAID-derived results by rejection rate for simulated data

	Rejection rate			
Variable	**0%**	**15%**	**30%**	**P-value**
Total matched shipments in 1 year period	532 ± 62	600 ± 65	610 ± 64	0.02
Unmatched islets in 1 year period				
Prop. among total produced	0.045 ± 0.030	0.078 ± 0.026	0.114 ± 0.025	<0.0001
Total IEQs (10^6)	0.68 ± 0.53	1.16 ± 0.53	1.67 ± 0.53	0.001
Prop. of requesters by % received islets #1				
0%	0.030 ± 0.020	0.035 ± 0.021	0.038 ± 0.023	0.73
1–49%	0.411 ± 0.068	0.331 ± 0.069	0.318 ± 0.045	0.004
50–99%	0.356 ± 0.054	0.376 ± 0.049	0.388 ± 0.029	0.31
100%	0.014 ± 0.012	0.028 ± 0.020	0.033 ± 0.021	0.06
>100%	0.189 ± 0.064	0.230 ± 0.055	0.224 ± 0.055	0.25
Prop. of requesters by % received islets #2				
0%	0.030 ± 0.020	0.035 ± 0.021	0.038 ± 0.023	0.73
1–49%	0.633 ± 0.079	0.543 ± 0.093	0.528 ± 0.079	0.02
50–99%	0.316 ± 0.064	0.396 ± 0.093	0.396 ± 0.064	0.03
100%	0.021 ± 0.022	0.026 ± 0.021	0.039 ± 0.029	0.27
>100%	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	---

The reported mean ± standard deviation statistics were based on 10 replicated 1-year simulations generated under islet supply versus demand ratio of 0.6. The % received islets “#1” and “#2” denote the requester-specific proportions of received IEQs among the requested minimum and ideal IEQs, respectively. P-values were obtained from one-way ANOVA comparing statistics under rejection rates of 0%, 15% and 30% in 10 replicated simulations. “---” = not applicable.