Automated Analysis of Microscopic Images of Isolated Pancreatic Islets

David Habart[1], Jan Švihlík[2.3], Jan Schier[5], Monika Cahová[4], Peter Girman[1], Klára Zacharovová[4], Zuzana Berková[4], Jan Kříž[1], Eva Fabryová[4], Lucie Kosinová[4], Zuzana Papáčková[4], Jan Kybic[2] and František Saudek[1].

AFFILIATIONS:

[1]Diabetes Center, Institute for Clinical and Experimental Medicine (IKEM), Prague, Czech Republic.

[2]Biomedical Imaging Algorithms Group, Department of Cybernetics, Faculty of Electrical Engineering, Czech Technical University in Prague, Czech Republic.

[3]University of Chemistry and Technology, Prague, Czech Republic.

[4]Center of Experimental Medicine, Institute for Clinical and Experimental Medicine (IKEM), Prague, Czech Republic.

[5]Department of Image Processing, Institute of Information Theory and Automation, The Czech Academy of Sciences, Prague, Czech Republic.

RUNNING HEAD: Automated Analysis of Islet Images

CORRESPONDENCE:

David Habart, MD, PhD
Diabetes Centre
Institute for Clinical and Experimental Medicine
Vídeňská 1958/9, PSČ 140 21, Prague 4, Czech Republic
Tel: +420 262 362 288; Fax: +420 261 362 820; Email: david.habart@ikem.cz.

ABSTRACT

Clinical islet transplantation programs rely on the capacities of individual centers to quantify isolated islets. Current computer-assisted methods require input from human operators. Here, we describe two machine learning algorithms for islet quantification, the trainable islet algorithm (TIA) and the non-trainable purity algorithm (NPA). These algorithms automatically segment pancreatic islets and exocrine tissue on microscopic images in order to count individual islets, and calculate islet volume and purity. References for islet counts and volumes were generated by the fully manual segmentation (FMS) method, which was validated against the internal DNA standard. References for islet purity were generated via the expert visual assessment (EVA) method, which was validated against the FMS method. The TIA is intended to automatically evaluate micrographs of isolated islets from future donors, after being trained on micrographs from a limited number of past donors. Its training ability was first evaluated on 46 images from four donors. The pixel-to-pixel comparison, binary statistics, and islet DNA concentration indicated that the TIA was successfully trained, regardless of the color differences of the original images. Next, the TIA trained on the four donors was validated on an additional 36 images from nine independent donors. The TIA was fast (67 sec/image), correlated very well with the FMS method ($R^2 = 1.00$ and 0.92 for islet volume and islet count, respectively), and had small REs (0.06 and 0.07 for islet volume and islet count, respectively). Validation of the NPA against the EVA method using 70 images from 12 donors revealed that the NPA had a reasonable speed (69 sec/image), an acceptable RE (0.14), and correlated well with the EVA method ($R^2 = 0.88$). Our results demonstrate that a fully automated analysis of clinical-grade micrographs of isolated

pancreatic islets is feasible. The algorithms described herein will be freely available as a

Fiji platform plugin.

KEY WORDS: islet transplantation, enumeration of islets, quality control, image

processing, image segmentation, machine-learning

INTRODUCTION

Since the breakthrough clinical trial published by the Edmonton group in 2000,

transplantation of isolated pancreatic islets has become a promising therapeutic approach

for patients with type 1 diabetes (24). Other centers have also reported good results with

respect to protection from severe hypoglycemic episodes, and even insulin independence

(2,13). However, these results are not globally consistent. Clinical islet transplantation

programs rely on the capacities of individual centers to evaluate the grafts they produce;

an inconsistent quantity of islets in the graft is a potential factor contributing to variation

in the success rate.

The methodology for the quantification of islets was first developed and standardized by

Ricordi (19). While many centers still count islets manually using a microscope and a 50

µm grid, over the past two decades the methodology has become computerized and

partially automated (4,7-10,12,17,18,27,31). Some centers have designed their own

macros using various semi-automated algorithms for digital image analysis, built for their

respective microscope systems (8,10). Other centers have developed dedicated software

applications for semi-automated digital image analysis using different, and sometimes

costly, platforms such as MetaMorph (9,15,17). However, none of these technologies is

fully automated, freely available, or widely used. To date, only one inter-laboratory study has been conducted, which found considerable inter-operator variation in islet count among 35 technicians using the grid-assisted manual count method. The study also showed reduced count variation among three technicians using a computer-assisted method (15). In addition to operator-related variation, Friberg et al. (8) found significant error associated with islet graft sampling. Sampling error can potentially be reduced by evaluating a sufficient number of samples. However, efficiency of current computer-assisted methods is limited by required manual input from operators. The only currently available fully automated analysis of digital images of islets is a dedicated islet cell counter (4). However, it is expensive and comparisons among operators and centers regarding count precision have yet to be undertaken. A common automated tool could improve good laboratory practice, especially for international clinical trials.

Clinically important islet graft parameters include the islet volume (measured in IE, islet equivalents), islet count, and islet size distribution. It has been shown (4,7-10,17,27,31) that this information can be extracted from islet micrographs using two-step image analysis. In the first step, the islets are distinguished from non-islet tissue via a process called 'segmentation'. Historically, the techniques used for segmentation have included manual thresholding (9) and watershed transformation (10). In the second step, areas of the segmented images attributed to islets are converted into individual islet volumes using a 3D islet model. Sphere (8,10,17,19) or prolate spheroid (9,17) models have traditionally been used for this purpose.

In this paper we evaluate the feasibility of a novel approach to the segmentation step of islet image analysis which becomes fully automated after an initial training period. We show that the trainable islet algorithm (TIA) can automatically segment a large number of previously unseen images in a short time, and with a high degree of accuracy, after been trained on a limited number of images. We also show that our non-trainable purity algorithm (NPA) can fully and automatically assess the purity of islets with an acceptable degree of accuracy.

MATERIALS AND METHODS

Study design

All work with animal and human tissue was performed in accordance with the Approval 712a/13 (G 13-04-10) issued by the Ethics Committee of the Institute for Clinical and Experimental Medicine and Thomayerova Hospital.

Original micrographs of isolated pancreatic islets were segmented, either manually or automatically, into islets and exocrine tissue in order to count the islets and calculate islet volume and purity. First, the manual methods used to generate the reference data (ground truth) were validated. Next, the automated methods were validated. The TIA for islet-counting was trained and validated using 82 original micrographs from 13 different donors (**Figure 1A**). The automated NPA, which calculates islet purity, was validated on a subset of 70 micrographs from 12 donors (**Figure 1B**).

Isolation of rat islets

Islets from Wistar rats (450 g, 5 months old) were harvested using a previously described method (22). Briefly, the main pancreatic duct was cannulated in an anesthetized animal. After sacrifice, the pancreas was filled with 15 mL of 1 mg/mL collagenase (Sevapharma, Czech Republic) before being excised and incubated for 22 min at 37ºC. Pure islets were then harvested from the layer of 1094 g/L of a discontinuous Ficoll gradient. Islets were stabilized overnight in 5 mL CMRL (Cellgro, MediaTech, VA, USA) medium supplemented with 10% Fetal Bovine Serum (FBS).

Isolation of human islets

Human islets were isolated from cadaver donors following a protocol based on that described in (20). Briefly, a mixture of collagenase (Serva, Germany or VitaCyte, IN, USA) and neutral protease (Serva, Germany) was instilled into the main duct of the pancreas. After incubation in a Ricordi chamber (20), the islets were separated from the exocrine tissue via centrifugation using a Cobe 2991 cell processor (Terumo BCT, Japan) and continuous Ficoll gradient (Biochrom, UK). The isolated islets were stabilized in a CMRL-based medium. Routine test samples were taken to assess graft quantity and quality.

Sample preparation, microscopy, and digital imaging

Isolated islets (100 µL) were stained in Petri dish with 10-25 µL of dithizone (DTZ). 50 mg of DTZ was dissolved in 800 µL of 70% ethanol with 400 µL 1 M NaOH, mixed and diluted with 15 mL phosphate-buffered saline (PBS) or Hanks' balanced salt solution (HBSS, Sigma-Aldrich, MO, US) or PBS supplemented with 0.2% albumin. After 1 min,

4 mL of HBSS/albumin was added and the dye dispersed. The islets were gently swirled towards the center. Samples were imaged using a zoom stereo microscope (model SZ60, Olympus, Japan) with inclined incident illumination on a white background, and a digital camera equipped with a CMOS 1/2" 3 MP image sensor (MicroCam, Bresser, Germany). Each RGB image was 2048 x 1536 pixels in size. The pixel size ranged from 2.4-4.9 µm for magnifications between 10 and 20x.

Determination of DNA content

After imaging, dithizone-stained pure islets were handpicked and washed, before being individually picked up in a total volume of 10-15 µL and placed into 150 µL of lysis buffer (1% SDS, 10mM EDTA and Proteinase K 50 µg/mL in 20 mM Tris-HCl, pH 7.4), sealed, and incubated overnight at 37ºC. DNA content in the vortexed lysate, diluted to at least 1:200, was determined using a Synergy-2 fluorometer (Ex 485/20 nm, Em 528/20 nm) (Bio-Tek, VT, USA). PicoGreen (1) fluorescent intercalating dye (P7589, Molecular Probes, OR, USA) was used as the probe because it is highly specific for dsDNA and suitable for cell lysates. Neither dithizone nor albumin would interfere with the assay at such a dilution. Lambda phage DNA was used to prepare the standard dilution curve (Molecular Probes, OR, USA).

The fully manual segmentation (FMS) method

Ground truth segmentations of islets were created using the fully manual segmentation method (FMS). An expert in islet counting used the GNU image manipulation program (GIMP 2.8.4) to manually delineate individual islets (25). Segmentation layers were

exported as png files and converted to 1-bit images. The best-quality standard segmentations were determined via the majority-vote method, based on the consensus of three independently created segmentations. Individual pixels were labeled as islets if they were marked as islets in at least two out of three sessions. Areas attributed to individual segmented islets were converted into normalized islet volumes, taking 1.767146 nL as 1 islet equivalent (IE) and using a table of IEs for discrete sets of islet diameters that increase in 50 µm increments (19). Islet diameter 'd' was calculated from individual islet area, assuming a circular shape where $d=2\sqrt{(area/\pi)}$ (8,15). Islets with a diameter smaller than 50 µm were not included in the assessment because they could not be reliably distinguished due to image resolution. Exocrine tissue was segmented using the same procedure described above. The total areas of exocrine tissue (Ae) and islets (Ai) were determined from each segmentation. Purity (Pu) was calculated as the ratio of the total area of the islets to the total tissue area: $Pu=Ai/(Ai+Ae)$.

Expert visual assessment (EVA) method

Five experienced operators were separately given images of dithizone-stained islets with varying purity. Each operator visually estimated purity as the fraction of islet tissue in the total tissue amount. The procedure involved two complementary steps. In the first step, the operator estimated the percentage of exocrine tissue in the total tissue and subtracted that percentage from 1.0 to obtain the percentage of remaining tissue. In the second step, the operator estimated the percentage of islets in the total tissue amount. The final estimated purity was calculated as average of the two steps. For each image, the average purity obtained from the five operators was used as a reference.

Automated algorithms

The TIA segments individual islets from an input image. To eliminate differences in illumination and camera settings, the algorithm first color-normalizes the images by assuming that the most common color is the background. The whole image is then color-transformed, such that the background becomes white and the dynamic range is fully used. Uneven illumination was corrected by fitting a second-order polynomial to approximately identified background pixels (29). The algorithm then uses a random forest classifier (3) to create a probability map of islets based on individual pixels, using RGB color components as features (30). A final binary classification (islets versus non-islets) is derived from the probability map by employing spatial regularization using GraphCut (16). The TIA calculates the volumes of the islets from the segmentations using the same model described for the FMS method.

For the assessment of islet purity by the NPA, binary segmentation is extended to a three-class segmentation (islets, exocrine tissue, background). We chose to use an unsupervised k-means clustering algorithm (14) for this task due to image variability, and a limited amount of available training data. After over-clustering the image pixels into seven color-based clusters, clusters corresponding to the three classes of interest are identified using two descriptors; the centroid ratio of the G and R channels (Gc/Rc) and the normalized cluster area (i.e., the area of every cluster divided by the maximum cluster area). Use of these descriptors was motivated by the observation that the largest image component is usually the background, and that islets and exocrine tissue are best distinguished by their

respective reddish and yellowish colors. Optimal descriptor value thresholds are determined empirically using a training set of images with known purities. In the final step, purity is calculated as the ratio of the islet area to the total area of the tissue.

The above-described algorithms were implemented for use on the Fiji platform (ImageJ 1.48e) (23) and can be requested for free from the authors.

Data evaluation and statistics

To compare the automatic and ground true segmentations, a pixelwise analysis method was developed using the statistical performance measures for the binary classification task. These measures included the number of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). The sensitivity (Se), specificity (Sp), precision (P), relative error (RE), and relative islet area (RIA) were calculated as:

$$Se=TP/(TP+FN)$$

$$Sp=TN/(TN+FP)$$

$$P=TP/(TP+FP)$$

$$RE=|(TP+FP)-(TP+FN)|/(TP+FP)$$

$$RIA=(TP+FP)/(TP+FN)$$

The performance of the trained TIA was evaluated via leave-one-out cross-validation (6).

Receiver operating characteristic (ROC) curves were generated using all data obtained from individual classifiers (in the cross-validation) (21).

Repeatability was assessed from coefficients of variation (CV). The results are expressed here as medians, 95% mean confidence intervals (CI), or by the mean ± standard deviation (SD). Means were compared using two-tailed unpaired Student's t-tests. Tukey box plots were used to visualize the data (26).

RESULTS

Validation of the FMS method

The repeatability of the FMS method was separately assessed on 135 pure rat islets and 1117 non-pure human islets (**Table 1**, sets A and B, respectively). Completely and partially embedded islets were present in about half of the images from set B (**Figure 2**, arrows). The results, summarized in **Table 2**, showed good repeatability of the FMS method with respect to islet area, count, and volume, irrespective of the presence or absence of exocrine tissue (CVs ranged from 0.02-0.05 for sets A and B). The limitations of the FMS method were tested by evaluating repeatability for a subclass of fully embedded islets from the same 12 images (set B, **Table 1**). The mean CV increased to 0.12 (95% CI: 0.04-0.20) when the triplicate FMS segmentations were evaluated exclusively for 35 fully embedded islets, indicating that such islets must not dominate images if high-precision ground truth data are required.

The accuracy of the FMS method was validated against the DNA standard. A total of 135 handpicked pure rat islets (**Table 1**, donor A) were distributed across 10 samples with increasing numbers of islets. The DNA content was determined for each sample (1), and islet volume was calculated using the FMS method. The correlation between DNA

content and the volume of islets present in the sample is shown in **Figure 3A** (R^2=0.98, with three independent volume estimates per image). The average DNA concentration was 8.4 ng/IE (95% CI: 7.8-9.0 ng/IE, n=30), which is in agreement with previously published values (11,12,18).

Validation of the EVA method

Applying the FMS method to exocrine tissue is possible, but extremely time-consuming. We therefore tested the robustness of a visual assessment of purity using the EVA method, based on the results of five experts. To validate this approach, the exocrine tissue was manually segmented by the FMS on a total of 36 images with varying islet purity, from three donors (**Table 1**, donors B-D). The mean CV for purity estimated by the five experts for the same images, was 0.18 (n=36). The EVA-based estimation of islet purity and the FMS-based purity determination were strongly correlated (R^2=0.95) (**Figure 3B**), and the slope of the regression line was close to one (1.01). The EVA method was therefore used to produce reference purities for NPA validation.

Training time of the TIA

To reduce the computational time required for TIA training, images and their corresponding ground truth segmentations were downsampled to 512 x 384 pixels. Receiver operating characteristic curves for the full size and downsampled images were generated for 12 images with various islet purities (donor B, **Table 1**) using leave-one-out cross-validation. The ROC curves generated for the full size and downsampled images were virtually identical (**Figure 4**), demonstrating that downsampling in training

does not lead to an appreciable loss of useful information. Downsampling reduced the average training time required on a standard laptop computer (Intel Core Duo CPU 2.54 GHz, 8 GB RAM) from 1200 s/image to 28 s/image. The segmentations created from full size and downsampled images were compared using pixel-wise measures of sensitivity (0.95±0.02), specificity (0.996±0.003), and precision (0.94±0.03). These measures were all close to one, indicating a very good agreement.

Trainability of the TIA

The TIA was developed to automatically segment clinical-grade islet images based on training data. Variation in image quality is expected in the clinical setting, despite standardized image acquisition. To establish the feasibility and robustness of our approach, we tested the hypothesis that training the TIA on images from several donors, with different acquisition set-ups, would impart the capacity to automatically segment similar images from multiple donors. Two training/testing schemes were designed to test this hypothesis.

In the first training scheme, the TIA was separately trained four times, each time on a single donor using all its image-segmentation pairs, producing trained tools TIA-A, TIA-B, TIA-C, and TIA-D. Each trained TIA was then applied to all of the images from the other three donors. Each trained TIA produced 32-36 automated segmentations, depending on the donor (donors A-D, **Table 1**). In total, 136 automatic segmentations were produced from 46 images. In the second training scheme, the TIA was trained 50 times; each time on a randomly selected half of the image-segmentation pairs from each

donor (**A-D, Table 1)**. Each trained TIA tool (TIA-R-No.) was then applied to the other

half of the images (i.e., the images not used for training the algorithm). The image-

segmentation pairs were randomized to avoid selection bias. As a result, individual

images were presented 16-36 times during 50 training iterations, and a similar number

(14-34) of times during testing. An example of a randomized allocation for one of the

trained TIA-R tools is shown in **Table 3**. In total, 1150 segmentations of 46 images were

produced. 1286 automated segmentations were compared in a pixel-wise manner to the

ground truth segmentations via direct visual comparisons (**Figure 5A**), and the relevant

statistics calculated. For the TIA trained on single donors, the RE of the islet area ranged

from 0.05-0.17, and the RIA ranged from 0.89-1.17 (**Figure 5B a-d**).

These data demonstrated that the image groups differed in appearance, and that the TIA

trained on a single group of images did not consistently produce good-quality

segmentations of images from other donors **(Figure 5B a-d)**. However, the mean RIA for

the TIA trained on half of images from all four donors was 0.99 and the RE was 0.07,

indicating a clear improvement in the quality of the segmentations **(Figure 5B r)**. Only

three images (out of 46), representing 51 observations (out of 1150), were incorrectly

segmented (outliers in **Figure 5B r**). The results indicate that, given a suitable training

set, the TIA is able to provide correct segmentations of images taken under different

conditions from multiple donors.

Repeatability and validation of the TIA against the DNA standard

The repeatability of the TIA from the above experiment was evaluated for 46 images and

50 randomizations. The mean CV was 0.02 (n=16-26). For three of the images (out of 46), the CV ranged from 0.05 to 0.15. These images were the above-mentioned outliers.

The accuracies of the 50 trained TIA-R tools described above were validated against the internal DNA standard in a manner similar to that undertaken for FMS validation, using the same set of images of pure islets (**Table 1**, donor A) and corresponding DNA samples. Here, each original image was automatically segmented approximately 25 times (range: 21-36 times, depending on the random allocation of each image to either the training or testing group). The resulting 250 randomized TIA-R segmentations were compared to the ground truth (for each of the 10 images, 14-29 of the 50 random assignments did not contain the image being evaluated). The RE of the islet area was 0.04 (95% CI: 0.03-0.05), and the correlation between the automatically estimated islet volumes and the corresponding DNA contents was R^2=0.99. The average islet DNA content was 8.6±1.7 ng/IE (95% CI: 7.6-9.6 ng/IE), which was in good agreement with the ground truth, and is within the previously published range (11,12,18).

Evaluation of the TIA on independent donors

The TIA is intended to be trained on a limited number of images from several donors, before being able to automatically analyze images from other donors not included in the training set (called here "independent donors"). This generalization ability of the TIA was tested here (the training ability of the algorithm tested above used different images, but the same donors for training and evaluation). The TIA was trained on 46 images from four donors (A-D, **Table 1**), producing a trained tool called TIA-ABCD. The trained tool

TIA-ABCD was then applied to 36 images of islets from nine independent donors (E1-E9, **Table 1, Figure 6**). The group of images was heterogeneous, created using two different stereo microscopes equipped with the same type of the camera but using various settings. The automatic segmentations, counts, and islet volumes were compared to ground truth data produced by the FMS method. The pixel-wise comparison revealed an acceptable RE (median RE=0.04, n=36). The automatically determined islet volumes and counts correlated well with the ground truth (R^2=1.0 and 0.92, for volumes and counts, respectively, n=36, **Figure 7A,B**). Finally, the automatically determined volumes and counts were expressed relative to the ground truth (TIA/FMS ratio) for each image (**Figure 7C**). Automatically determined islet volumes were marginally overestimated by the TIA-ABCD (median ratio 1.03, n=36), and islet counts were the same as the ground truth (median ratio 1.00, n=36). **Figure 7B** shows very good agreement between the two methods in the number of islets for all images, with the exception of the two images with the greatest numbers of islets. In these cases it was sometimes difficult to distinguish one islet from two closely spaced smaller ones. Exocrine tissue was also occasionally erroneously classified as islets. The automatic determination of islet parameters for all 36 images took 40 min (mean: 67 s/image).

Validation of the NPA for automatic purity assessment

The NPA was designed to automatically estimate islet graft purity from micrographs. The algorithm does not require training; the parameters used to identify the classes of interest (see the section on automated algorithms in the Methods) were determined using 12 islet images of varying purity (**Table 1**, donor B). The FMS segmentations with three classes

(islets, exocrine tissue and background) were used as reference data. The NPA was validated on 70 images from the remaining 12 donors (A,C,D,E1-E9, **Table 1**). For validation, five experts estimated purities using the EVA method. The degree of variability among the experts was similar to that observed during validation of the EVA method (CV=0.19, n=70). The automatically determined and reference purities were strongly correlated (R^2=0.88), and the slope of the regression line was close to one (**Figure 8A**). The median relative NPA purity for all 70 images was close to one (1.01) and the mean purity was 1.11. A Tukey box plot identified 7 image outliers (out of 70 images) (**Figure 8B**). The automatic purity assessment of all 70 images was accomplished within 80 min (mean: 69 s/image).

DISCUSSION

In this work we developed, implemented, and validated two algorithms (i.e., the TIA and the NPA) designed to automatically segment pancreatic islets and exocrine tissue from microscopic images for the purposes of enumerating individual islets, estimating islet volumes, and islet purities.

The islet count and volume are major determining factors for the transplantability of grafts to human recipients. Another factor that could potentially restrict graft transplantability is the degree of contamination by the exocrine tissue. In this case the limiting factor is not the purity itself, but the total volume of the graft tissue in mL; the maximum allowable volume is a multiple of the necessary volume of islets. Therefore, while the highest possible purity is desirable, its precise determination is not essential. Instead, the ability to categorize samples as having a high, intermediate, or low purity is

acceptable. Supervised training of a machine learning algorithm can greatly improve the precision of automated image segmentation. However, training requires the creation of reference segmentations, which is time-consuming and apparently not necessary for purity assessment. These facts led us to develop two different automatic tools, the TIA and the NPA. Our strategy involved using a trainable algorithm for automated islet segmentation (TIA), and an unsupervised algorithm for segmentation of the exocrine tissue (NPA).

The TIA and NPA tools were developed as plugins for the free, professional image analysis software program ImageJ. High-quality ground truth segmentations are necessary to employ machine learning, and the methods used for their generation need to be validated. We consider the manual delineation of islets (i.e., the FMS method) to be a preferable approach for creating ground truth segmentations because it allows experts to treat islets individually, despite variation in the color composition of islets on the same image. We chose two objectively measurable parameters for validation of the FMS method; islet DNA content, and the repeatabilities of islet count and volume determination. The first parameter is based on a sensitive and robust DNA measurement and the determination of islet volume by the validated method. Three other independent assessments have been previously reported, ranging from 4 and 10 ng/IE (11,12,18). The FMS method yielded a high degree of correlation between islet volume and DNA content (**Figure 3A**), and was in good agreement with previously reported absolute values (8.4±1.6 ng/IE). Pure rat islets (not human) were used for two reasons; first, to avoid DNA degradation (rat islets have easier logistics and are not so easily damaged by extensive handling), and second, to minimize the likelihood of false overestimation due

to interference by non-segmented exocrine tissue with its own DNA. The images of rat

islets were acquired at a similar magnification to the clinical-grade images (**Table 1**). The

borders of the rat islets were visually similar to those of human islets, for the purpose of

manual delineation. That the similarity in the islet images from the two species was

sufficient was also supported by the similar repeatabilities obtained for pure rat islets and

non-pure human islets (sets A and B, **Table 2**). The second parameter, the repeatability of

the FMS method was high (**Table 2**), even for images containing a small number of

embedded islets.

The training ability of the TIA tool was validated in two steps. The first step used 46

images from donors included in the training of 50 TIA tools (**Figure 5B, r**). These 50

trained TIA tools demonstrated that, provided that a sufficient qualitative range of images

is used for training, the algorithm is capable of segmenting islet images from various

donors with high precision (islet DNA content was 8.6±1.7 ng/IE), good repeatability

(CV=0.02), and a low RE (0.07). In the second step, the application of the TIA to future

images was modeled using 36 images from nine additional donors that were not included

in the training set (**Figure 6**). Based on our experience with the FMS method, we

excluded images unsuitable for manual analysis. Excluded images had borders that were

difficult to define by the expert producing the manual segmentation: islets substantially

covered by exocrine tissue, small islets embedded within large pieces of exocrine tissue

(**Figure 9A**), large and thick chunks of exocrine tissue (**Figure 9A**), and grossly over-

digested islets with many small fragments or blurred borders (**Figure 9B**). Micrographs

containing unevenly stained islets with pale interiors (**Figure 9C**), or very weakly stained

islets (**Figure 9D**) were also problematic. Furthermore, images of islets cultured for several days were excluded due to the gradual, uneven, and patchy loss of the dithizone staining, which made the reliable delineation of islets very difficult. However, images with very different color compositions (**Figure 6**), such as a greenish or bluish background or purple islets, were included in the analysis as long as the FMS method could be applied with confidence. These color variations can be likely attributed to incorrect camera settings. Given these limitations, the TIA trained on four donors (TIA-ABCD) and applied to nine other donors correlated well with the ground truth in terms of islet volume and count (**Figure 7A,B**). The mean and median relative islet volume and count were close to one (**Figure 7C**).

The TIA is trained once on a range of images at a reasonable training speed (two images per minute). Following training, the algorithm can be used to analyze future images at a pace of one image per minute and with an accuracy that makes it useful in a clinical setting. At present, images with multiple islets in spatial contact with one another should be avoided because they risk being segmented as a single unit, leading to an overestimation of islet volume and an underestimation of islet count. Future upgrades of the TIA tool currently under development may be able to automatically resolve the separation of closely spaced islets (28). Another area of future study is the complicated relationship between islet shape in 2D images and islet volume; here we simply apply the currently used classical spherical islet shape model (19). An effort to obtain more accurate automated estimates of the volumes of islets from their 2D images is underway (5), to complement the automated segmentation process described in this work.

The manual delineation of exocrine tissue by the FMS method is laborious; therefore, for this study a substitute method was needed. The data obtained by the FMS method were used as a standard to validate a less work-intensive approach, the EVA method. The experts compared reasonably well among themselves, with a CV of 0.19, (which was expected for a visual assessment), and with the FMS method (**Figure 3B**). Given the less stringent requirement for purity assessment (i.e., the use of purity categories rather than percentages), the EVA method with a CV<0.20 was considered acceptable for creating the reference dataset for NPA validation. The NPA compared well with this reference dataset (**Figure 8**). However, the spread of the observed values was expectedly wider, compared to the TIA. The performance of the NPA (one image per minute) is comparable to that observed for the EVA method and can be considered for clinical practice. In future work it would be interesting to compare our TIA and NPA tools to already established computer-assisted methods (8,10,17). Currently, it is not possible for us to undertake such a comparison because those methods are not readily available to us.

We conclude that it is feasible to combine a trainable algorithm for the high-precision determination of islet parameters with an unsupervised algorithm for the assessment of graft purity, for a fully automated assessment of clinical-grade images of isolated pancreatic islets. Our tools provide an objective and rapid way of analyzing multiple islet microscopy images before transplantation. Currently, we are actively looking for collaborating institutions to assist in improving these tools further by assembling a larger dataset of qualitatively diverse images from different laboratories.

DISCLOSURE

The authors of this manuscript have no conflicts of interest to disclose as described by the

Cell Transplantation.

FIGURE LEGENDS

Figure 1: *Study design.* The TIA and NPA, two algorithms for the automated assessment of islet images, are shown in the black boxes. The TIA (A) was trained using the original islet images and the corresponding ground truth segmentations, produced by the FMS method. The training set consisted of a number of image-segmentation pairs. The dashed arrow represents downsampling of the images during training. The TIA was validated using image-segmentation pairs (i.e., the testing set) composed of images not used for training. The individual islets were counted and their volumes calculated from the TIA and the FMS method, and compared for TIA validation. The NPA (B) was not trained. This method of purity assessment was validated against results generated by the EVA method.

Figure 2: *Images and their segmentations.* Dithizone-stained islet graft images (A-D) representing images from four donors (A-D, **Table 1**). Arrows (B-D) indicate fully embedded islets. Panel (E) shows the ground truth segmentation of image (C). In (F), the same image was segmented automatically by the trained TIA.

Figure 3: *Validation of the FMS and EVA methods.* (A) The FMS was validated against an internal DNA standard. Ten samples of 1-58 pure rat islets were stained with dithizone, and microscope images of the samples were acquired and segmented in triplicate by the FMS method. The DNA content of each sample was determined using the PicoGreen method. Islet volumes (IE) were calculated from the triplicate

segmentations. The extent of the correlation between DNA content and islet volume was assessed by linear regression. (B) The EVA method was validated against the validated FMS method. Five independent experts visually assessed the purity of 46 islet images. The mean expert-determined purities (y-axis) were calculated and correlated with the purities determined by the FMS method (x-axis).

Figure 4: *Image downsampling for rapid training.* The TIA was trained on full size and downsampled image-segmentation pairs. The possible loss of useful information was assessed by the leave-one-out cross-validation method on 12 images from set B, **Table 1**. Receiver operating characteristic (ROC) curves were constructed for three image sizes and compared. Circle: full size images, cross: half-size images, X: quarter-size images. The false positive rate (FPR=1-specificity) is on the x-axis. The true positive rate (TPR=sensitivity) is on the y-axis.

Figure 5: *Training capability of TIA.* Training capability was assessed via a pixel-to-pixel comparison between the automatically determined and ground truth segmentations. (A) Visual output of the pixel-wise analysis for a single image (**Figure 2B**): true negative (blue), true positive (red), false negative (green), and false positive (white) pixels are visualized. Islets segmented by the FMS method are comprised of red and green pixels. Islets segmented by the TIA method are comprised of red and white pixels. The long and short arrows indicate fully and partially embedded islets, respectively. (B) The TIA was trained, either on images from a single donor (a-d, x-axis), or on a randomly selected half of each donor's images (50 randomized allocations, r, x-axis). Trained tools TIA-A, -B, -

C, and -D were tested on 34, 36, 32, and 34 images from three (out of 4) donors not used

for training (a-d, x-axis). The 50 trained TIA-R were tested on a total of 46 images from

all four donors, where each image was used 14-34 times, producing 1150 segmentations

(r, x-axis). The Tukey box plot depicts the pixels attributed to islets by the TIA, relative

to the pixels assigned to islets by the FMS method (y-axis); median (horizontal line), 1st

and 3rd quartiles (box), 1.5 interquartile range (whiskers).

Figure 6: *Images from the independent donors.* Eight images (E1-E9) representing eight

of the nine donors of 36 images used for validation of the TIA-ABCD tool. Details of

these images are given in **Table 1**.

Figure 7: *Validation of the TIA on independent donors.* The TIA-ABCD tool was trained

on 46 images from four donors (A-D, **Table 1**), and tested on 36 images from nine other

donors (E1-E9, **Table 1**, **Figure 6**), which were ground truthed via the FMS method. (A)

The correlation between the values for islet volume obtained by the TIA and the FMS

method. (B) The correlation between the values for islet count obtained by the TIA and

the FMS method. The Tukey box plot in (C) shows TIA-ABCD volumes and counts

relative to the ground truth (y-axis); mean (cross), median (horizontal line), 1st and 3rd

quartiles (box), 1.5 interquartile range (whiskers).

Figure 8: *Automatic purity assessment by the NPA.* Islet purities were assessed by the

NPA for 70 islet micrographs from 12 donors, and compared to reference purities

produced via the EVA method. (A) Correlation between automatically estimated islet

purities (y-axis, purities determined via the NPA) and the reference purities (x-axis,

purities determined via the EVA method). (B) The Tukey box plot shows the relative

NPA purities for 70 donors; mean (cross), median (central line), 1st and 3rd quartiles

(box), 1.5 interquartile range (whiskers).

Figure 9: *Examples of images unsuitable for the TIA.* Unsuitable images contain (A)

large chunks of exocrine tissue (arrow) or small islets embedded within large pieces of

exocrine tissue (arrowheads), (B) over-digested islets with blurred borders, (C) poorly

stained islets with pale centers, and (D) weakly stained islets.

REFERENCES:

1. Ahn, S. J.; Costa, J.; Emanuel, J. R. PicoGreen quantitation of DNA: effective evaluation of samples pre- or post-PCR. Nucleic Acids Research. 24: 2623–2625; 1996.

2. Alejandro, R.; Barton, F. B.; Hering, B. J.; Wease, S.; Collaborative Islet Transplant Registry Investigators. 2008 Update from the Collaborative Islet Transplant Registry. Transplantation. 86: 1783–1788; 2008.

3. Breiman, L. Random forests. Machine Learning. 45: 5–32; 2001.

4. Buchwald, P.; Berman, D. M.; Szust, J.; Robassa, A.; Fraker, C.; Echeverri, F.; Bernal, A.; Inverardi, L.; Kenyon, N. S.; Ricordi, C. Assessment of the accuracy and reproducibility of a digital image analysis based on Automatic islet cell counter (ICC) using microsphere mixtures as well as human and NHP islets. 2011 Joint congress of Cell transplant society and International Xenotransplantation Assosiation. 1–1; 2011.

5. Dvorak, J.; Svihlik, J.; Habart, D.; Kybic, J. Comparison of volume estimation methods for pancreatic islet cells. In: Gimi, B.; Krol, A., ed. SPIE; 2016: vol. 9788. 97881K–97881K–6.

6. Evgeniou, T.; Pontil, M.; Elisseeff, A. Leave one out error, stability, and generalization of voting combinations of classifiers. Machine Learning. 55: 71–97; 2004.

7. Fetterhoff, T. J.; Wile, K. J.; Coffing, T. J.; Cavanagh, T. J.; Wright, M. J. Quantitation of isolated pancreatic islets using imaging technology. Transplant Proc. 26: 3351; 1994.

8. Friberg, A. S.; Brandhorst, H.; Buchwald, P.; Goto, M.; Ricordi, C.; Brandhorst, D.; Korsgren, O. Quantification of the islet product: presentation of a standardized current good manufacturing practices compliant system with minimal variability. Transplantation. 91: 677–683; 2011.

9. Girman, P.; Kriz, J.; Friedmansky, J.; Saudek, F. Digital imaging as a possible approach in evaluation of islet yield. Cell transplantation. 12: 129–133; 2003.

10. Gmyr, V.; Bonner, C.; Lukowiak, B.; Pawlowski, V.; Dellaleau, N.; Belaich, S.; Aluka, I.; Moermann, E.; Thevenet, J.; Ezzouaoui, R.; Queniat, G.; Pattou, F.; Kerr-Conte, J.

Automated Digital Image Analysis of Islet Cell Mass Using Nikon's Inverted Eclipse Ti Microscope and Software to Improve Engraftment May Help to Advance the Therapeutic Efficacy and Accessibility of Islet Transplantation Across Centers. Cell transplantation. 24: 1–9; 2015.

11. Hopcroft, D. W.; Mason, D. R.; Scott, R. S. Standardization of insulin secretion from pancreatic islets: validation of a DNA assay. Horm. Metab. Res. 17: 559–561; 1985.

12. Huang, H.-H.; Ramachandran, K.; Stehno-Bittel, L. A replacement for islet equivalents with improved reliability and validity. Acta Diabetol. 2012.

13. Investigators, C. 2011 Update from the Collaborative Islet Transplant Registry (CITR). Am J Transplant. 12: 161–161; 2012.

14. Kanungo, T.; Mount, D. M.; Netanyahu, N. S.; Piatko, C. D.; Silverman, R.; Wu, A. Y. An efficient k-means clustering algorithm: analysis and implementation. IEEE Transactions on Pattern Analysis and Machine Intelligence. 24: 881–892; 2002.

15. Kissler, H. J.; Niland, J. C.; Olack, B.; Ricordi, C.; Hering, B. J.; Naji, A.; Kandeel, F.; Oberholzer, J.; Fernandez, L.; Contreras, J.; Stiller, T.; Sowinski, J.; Kaufman, D. B. Validation of methodologies for quantifying isolated human islets: an Islet Cell Resources study. Clinical transplantation. 24: 236–242; 2010.

16. Kolmogorov, V.; Zabih, R. What energy functions can be minimized via graph cuts. IEEE Transactions on Pattern Analysis and Machine Intelligence. 26: 147–159; 2004.

17. Niclauss, N.; Sgroi, A.; Morel, P.; Baertschiger, R.; Armanet, M.; Wojtusciszyn, A.; Parnaud, G.; Muller, Y.; Berney, T.; Bosco, D. Computer-assisted digital image analysis to quantify the mass and purity of isolated human islets before transplantation. Transplantation. 86: 1603–1609; 2008.

18. Pisania, A.; Papas, K. K.; Powers, D. E.; Rappel, M. J.; Omer, A.; Bonner-Weir, S.; Weir, G. C.; Colton, C. K. Enumeration of islets by nuclei counting and light microscopic analysis. Lab Invest. 90: 1676–1686; 2010.

19. Ricordi, C.; Gray, D. W.; Hering, B. J.; Kaufman, D. B.; Warnock, G. L.; Kneteman, N. M.; Lake, S. P.; London, N. J.; Socci, C.; Alejandro, R. Islet isolation assessment in man and large animals. Acta Diabetol Lat. 27: 185–195; 1990.

20. Ricordi, C.; Lacy, P. E.; Scharp, D. W. Automated islet isolation from human pancreas. Diabetes. 38 Suppl 1: 140–142; 1989.

21. Rodríguez, J. D.; Pérez, A.; Lozano, J. A. Pattern Recognition. Pattern Recognition. 46: 855–864; 2013.

22. Saudek, F.; Cihalova, E.; Karasova, L.; Kobylka, P.; Lomsky, R. Increased glucagon-stimulated insulin secretion of cryopreserved rat islets transplanted into nude mice. J Mol Med. 77: 107–110; 1999.

23. Schindelin, J.; Arganda-Carreras, I.; Frise, E.; Kaynig, V.; Longair, M.; Pietzsch, T.; Preibisch, S.; Rueden, C.; Saalfeld, S.; Schmid, B.; Tinevez, J.-Y.; White, D. J.; Hartenstein, V.; Eliceiri, K.; Tomancak, P.; Cardona, A. Fiji: an open-source platform for biological-image analysis. Nat Methods. 9: 676–682; 2012.

24. Shapiro, A. M.; Lakey, J. R.; Ryan, E. A.; Korbutt, G. S.; Toth, E.; Warnock, G. L.; Kneteman, N. M.; Rajotte, R. V. Islet transplantation in seven patients with type 1 diabetes mellitus using a glucocorticoid-free immunosuppressive regimen. N Engl J Med. 343: 230–238; 2000.

25. Solomon, R. W. Free and open source software for the manipulation of digital images. AJR Am J Roentgenol. 192: W330–4; 2009. www.gimp.org, 1997-2014, retrieved on Feb 8, 2013.

26. Spitzer, M.; Wildenhain, J.; Rappsilber, J.; Tyers, M. BoxPlotR: a web tool for generation of box plots. Nat Methods. 11: 121–122; 2014.

27. Stegemann, J. P.; O'Neil, J. J.; Nicholson, D. T.; Mullon, C. J. Improved assessment of isolated islet tissue volume using digital image analysis. Cell transplantation. 7: 469–478; 1998.

28. Svihlik, J.; Kybic, J.; Habart, D. Automated separation of merged Langerhans islets. SPIE 2016. In M. A. Styner & E. D. Angelini (Eds.), (Vol. 9784, p. 978438).

29. Svihlik, J.; Kybic, J.; Habart, D. Color normalization for robust evaluation of microscopy images. SPIE, Medical Imaging 2016. 9599; 2015.

30. Svihlik, J.; Kybic, J.; Habart, D.; Berkova, Z.; Girman, P.; Kriz, J.; Zacharovova, K. Classification of microscopy images of Langerhans islets. 90341S–90341S–8; 2014.

31. Wile, K. J.; Fetterhoff, T. J.; (null); Cavanagh, T. J.; Wright, M. J. Morphologic analysis of pancreatic islets automated image analysis. Transplant Proc. 26: 3441; 1994.
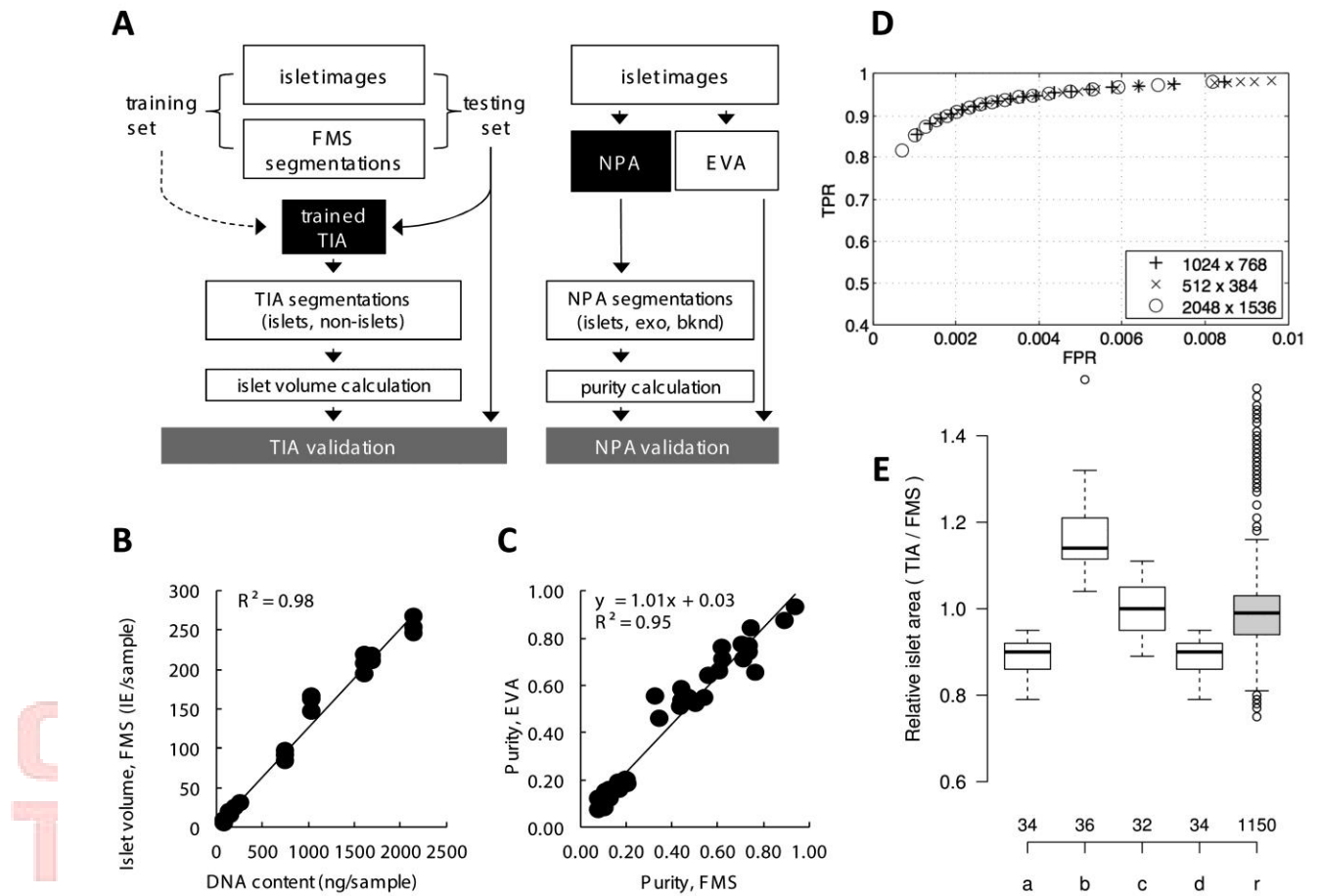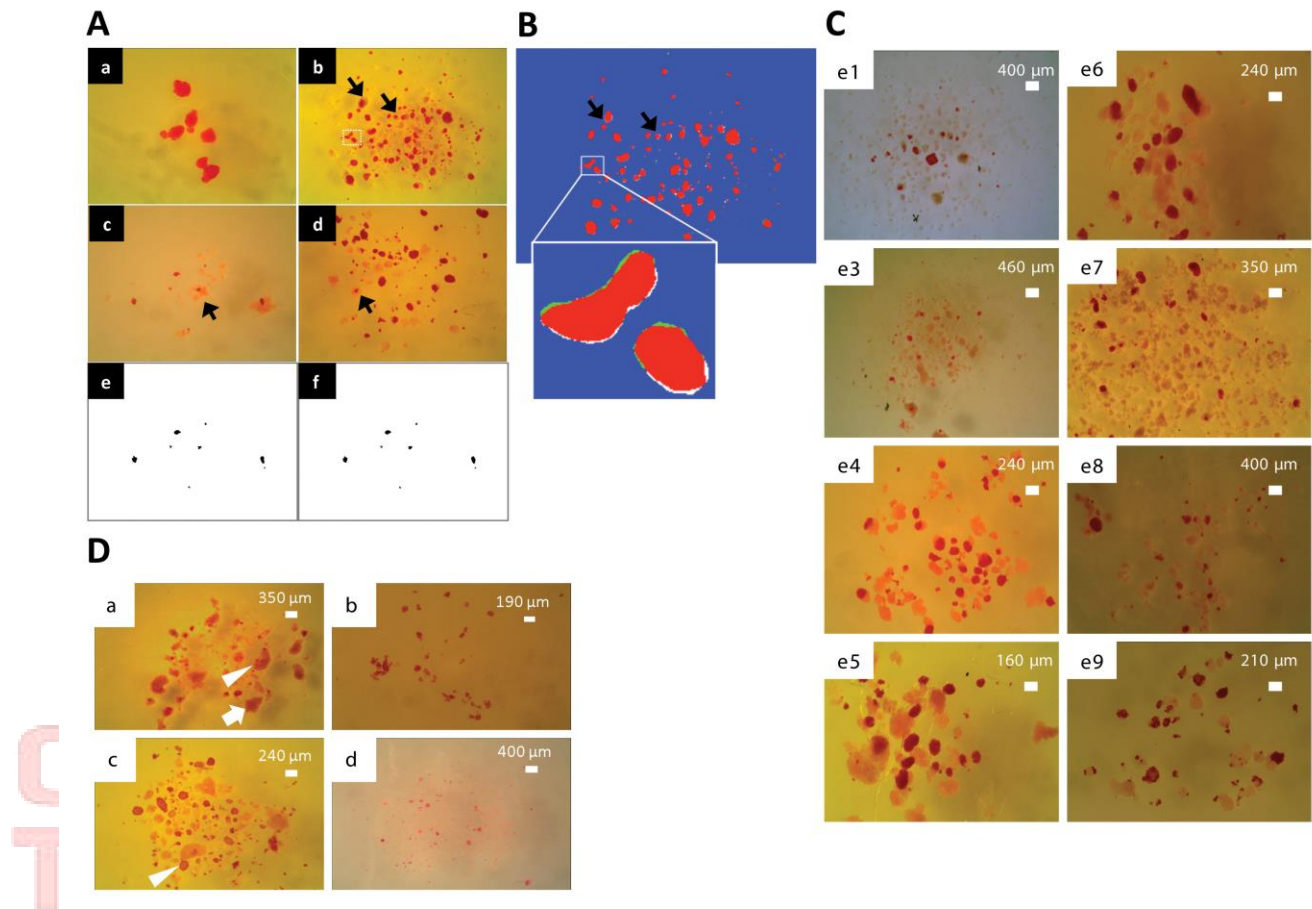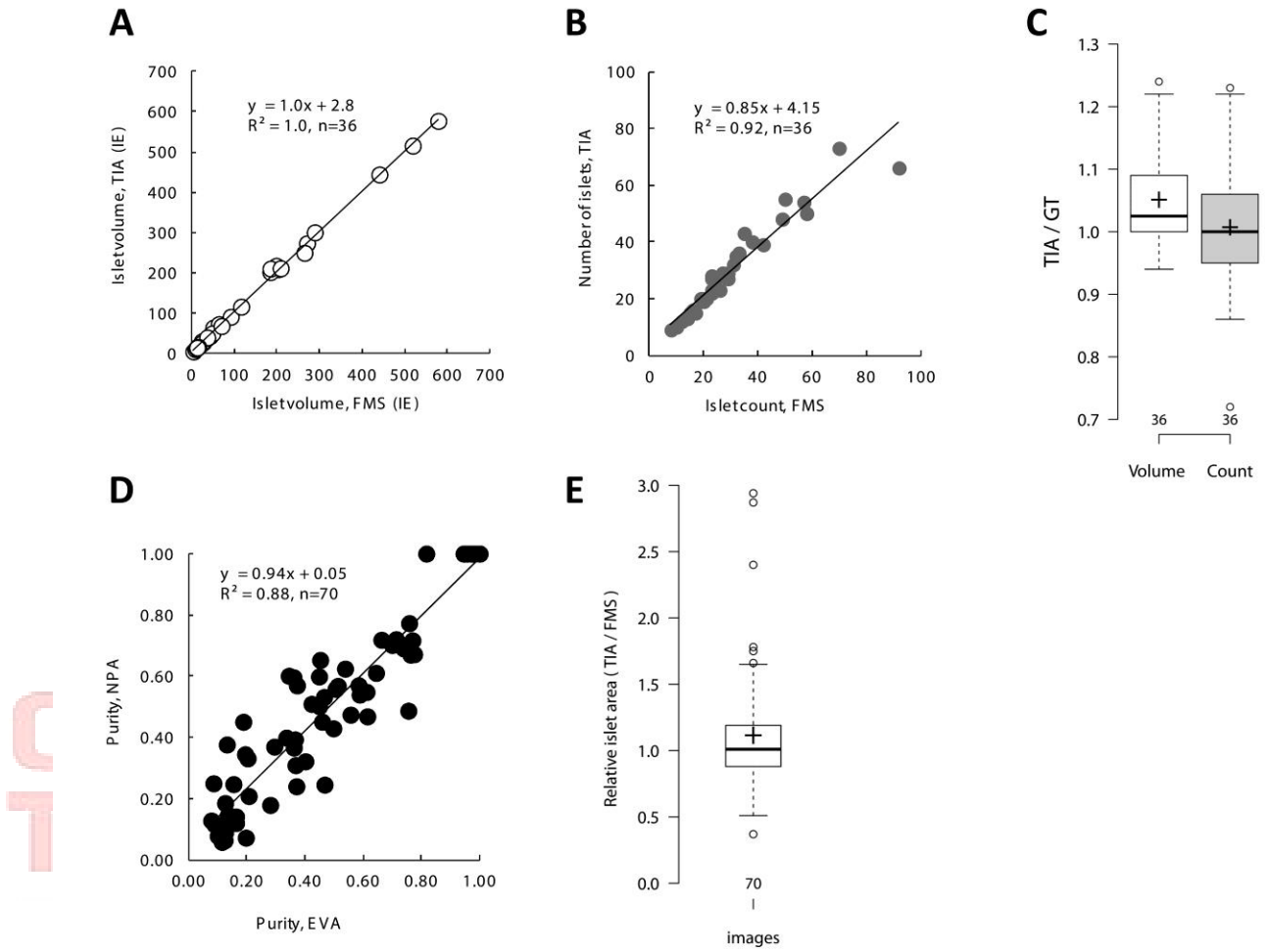
Figure 1

Figure 2

Figure 3

Table 1: The image-segmentation pairs used in the study

Donor	Images					
	No. of images	Magni-ficanion	No. of islets per image	Islet size (µm)	Median purity (%)	Species
A	10	20x	1-58	50-600	99	rat
B	12	10x,15x	63-149	50-600	75	human
C	14	15x	4-13	50-250	15	human
D	10	20x	24-67	50-300	65	human
E1	4	10x	16-38	50-450	62	human
E2	4	10x	12-20	50-300	37	human
E3	3	10x	10-23	50-350	12	human
E4	4	20x	29-70	50-550	46	human
E5	4	20x,30x	21-57	50-550	40	human
E6	4	20x	24-58	50-550	37	human
E7	5	15x	8-27	50-350	13	human

E8	4	12x	17-31	50-500	36	human
E9	4	25x	20-92	50-250	60	human

* high quality reference segmentation by the majority vote method

Table 2: Repeatability of the FMS method

Set	Experts' coefficient of variation			No. of islets	n
	Islet area median (95% CI)	Islet count median (95% CI)	Islet volume median (95% CI)		
A	0.02 (0.01 - 0.03)	0.02 (0.01 - 0.07)	0.05 (0.01 - 0.13)	135	30
B	0.03 (0.03 - 0.04)	0.03 (0.02 - 0.04)	0.03 (0.03 - 0.05)	1117	36

n, number of standard segmentations

Table 3: Trained tool TIA-R-21/50

| Donor (Image ID) | Images and Ground truth segmentations | |
	Training group image ID	Testing group image ID
A (1-10)	2,5,7,8,9,10	1,3,4,6,11,12
B (11-22)	15,17,18,19,21	13,14,16,20,22
C (22-36)	24,25,26,27,29,33,34	23,28,30,31,32,35,36
D (37-46)	38,39,42,43,44	37,40,41,45,46
total 46	random 23	remaining 23

ID, identification code.