



# HHS Public Access

Author manuscript

*Multimodal Learn Clin Decis Support Clin Image Based Proc (2020)*. Author manuscript; available in PMC 2021 June 09.

Published in final edited form as:

*Multimodal Learn Clin Decis Support Clin Image Based Proc (2020)*. 2020 October ; 12445: 13–23.

doi:10.1007/978-3-030-60946-7\_2.

## Prediction of Type II Diabetes Onset with Computed Tomography and Electronic Medical Records

Yucheng Tang<sup>1</sup>, Riqiang Gao<sup>1</sup>, Ho Hin Lee<sup>1</sup>, Quinn Stanton Wells<sup>2</sup>, Ashley Spann<sup>2</sup>, James G. Terry<sup>2</sup>, John J. Carr<sup>2</sup>, Yuankai Huo<sup>1</sup>, Shunxing Bao<sup>1</sup>, Bennett A. Landman<sup>1,2</sup>

<sup>1</sup>Vanderbilt University, , Nashville, USA

<sup>2</sup>Vanderbilt University Medical Center, , Nashville, USA

### Abstract

Type II diabetes mellitus (T2DM) is a significant public health concern with multiple known risk factors (*e.g.*, body mass index (BMI), body fat distribution, glucose levels). Improved prediction or prognosis would enable earlier intervention before possibly irreversible damage has occurred. Meanwhile, abdominal computed tomography (CT) is a relatively common imaging technique. Herein, we explore secondary use of the CT imaging data to refine the risk profile of future diagnosis of T2DM. In this work, we delineate quantitative information and imaging slices of patient history to predict onset T2DM retrieved from ICD-9 codes at least one year in the future. Furthermore, we investigate the role of five different types of electronic medical records (EMR), specifically 1) demographics; 2) pancreas volume; 3) visceral/subcutaneous fat volumes in L2 region of interest; 4) abdominal body fat distribution and 5) glucose lab tests in prediction. Next, we build a deep neural network to predict onset T2DM with pancreas imaging slices. Finally, motivated by multi-modal machine learning, we construct a merged framework to combine CT imaging slices with EMR information to refine the prediction. We empirically demonstrate our proposed joint analysis involving images and EMR leads to 4.25% and 6.93% AUC increase in predicting T2DM compared with only using images or EMR. In this study, we used case-control dataset of 997 subjects with CT scans and contextual EMR scores. To the best of our knowledge, this is the first work to show the ability to prognose T2DM using the patients' contextual and imaging history. We believe this study has promising potential for heterogeneous data analysis and multi-modal medical applications.

### Keywords

Type II diabetes; Electronic medical records; Computed tomography; Metabolic syndrome; Disease onset prediction

## 1 Introduction

Type II diabetes mellitus (T2DM) [1, 2, 3] is a common and significant chronic disease with both inherent and environmental causes. T2DM is characterized by obesity with attendant risk factors including hyperglycemia, hypertension, and hyperglycemia stemming from insulin resistance [4, 5, 6]. Potential markers of T2DM include the aforementioned risk factors as well as regional obesity and pancreas changes which can be learned from patients' imaging and diagnostic history [7]. Clinical framing of these variables relative to T2DM are complex through multiple risk factors, *e.g.*, body mass index (BMI), pancreas tissue volume, visceral/subcutaneous fat distribution, and glucose tests. Previous works have shown these hand-crafted features can be used to classify the presence of T2DM [8, 9]. However, the longer-term effects of risk factors are less well understood.

Clinical evaluation of patients with potential risk is performed by examining their electronic medical records (EMR) including 1) demographics, 2) ICD-9 codes, 3) lab tests, or 4) clinical and medication histories. With the advent of EMR, researchers have used machine learning and data mining methods in diabetes research, such as predictive biomarker identification, disease prediction, and diagnosis [10]. Mani et al. used demographic, clinical and lab parameters from EMR with different machine learning algorithms (linear based, one sample-based, decision tree based, and one kernel-based classifiers) for predicting T2DM risks in the period six months to one-year prior to diagnosis of diabetes [8]. Zheng et al. proposed a framework for identifying subjects with and without T2DM from EMR via feature engineering and similar machine learning methods [11]. Anderson et al. employed logistic regression and a random-forests probabilistic model on patients' full EMR or restricted EMR and showed EMR phenotyping can out-perform predict diagnosis of T2DM with conventional screening methods [7]. Meanwhile, there are few researches applying medical imaging studies (MRI, abdominal CT) to understand the association between T2DM and tissue composition as well as volumetric measurements. For instance, livers volumes, size of pancreas and body fat content [12, 13] are related to T2DM.

Recent advances, such as use of multi-modal machine learning [14, 15], bring opportunities for medical applications derived from both EMR and medical image data [16, 17]. Virostko et al. discovered pancreas volume declines with disease duration in type I diabetes (T1D) patients by using electronic medical record and magnetic resonance imaging (MRI)/ CT scans [9]. Chaganti et al. developed a method for multi-modal big data studies in MRI image processing that used EMR information, to classify diabetes patients in orbit CT [18]. However, the above works do not focus on early prognosis. Since the abdominal CT is becoming a routinely acquired imaging technique [19, 20], our goal is to explore the feasibility of combining EMR and CT to predict risk factors for T2DM a year prior to diagnosis.

In this work, we investigate five different types of EMR features to predict the onset of T2DM. We extract the 1) patient demographics, 2) pancreas volume, 3) fat volume in L2 region of interests (ROI), 4) visceral/subcutaneous fat distribution along the abdomen, and 5) glucose lab tests from each patient's clinical history. We perform each configuration of EMR

in an ablation scheme for T2DM onset prediction. For each subject, we formulate an EMR feature vector describing their clinical history from each configuration. We show that each contextual feature from patients' clinical history improves the prediction of onset T2DM. Next, we construct a deep neural network for encoding pancreas CT imaging slices for the T2DM onset prediction. Inspired by previous works on EMR-guided image processing [8], we developed a framework that combines CT images and EMR features.

We conduct experiments on 997 subjects using the “case-control” design [21, 22]. 401 subjects are with diagnoses of T2DM. The remaining 596 subjects are in the control group of non-diabetes. For extracting anatomic volumes, we trained the segmentation model for this work using 3D U-net with 100 CT scans and labels from multi-atlas labeling MICCAI challenge 2015 [23]. We believe this work could motivate further investigation of heterogenous data analysis and EMR guided multimodal medical applications.

In summary, our contributions in this work are: 1) We present a “case-control” study design for onset prediction of T2DM; 2) We evaluate that five different configurations of EMR features can contribute to prediction of onset T2DM a year prior to diagnosis; and 3) We show that EMR-image multi-modal framework for heterogenous data analysis improves predictive power for medical applications.

## 2 Method

Our proposed method involves five different categories of EMR features and a deep neural network architecture for encoding pancreas imaging slices, as illustrated in Fig. 1.

### 2.1 Data of T2DM Studies

A total of 997 de-identified subjects were selected and retrieved from our medical center under the institutional review board (IRB) approval from 6317 studies with diagnosis codes involving spleen abnormalities (cohort A). non-spleen abnormalities (cohort B). The dataset follows “case-control” ICD-9 code design [22]: (1). Case: T2DM subjects identified from ICD-9 code with diagnostic date (*ICD-9 = 250.## group of type 2 diabetes diagnosis*); (2) Control: non-diabetes subjects without diabetes ICD-9 codes or medication. There are 401 cases of T2DM diagnosis and 596 subjects under the control group of non-diabetes who were chosen for having similar imaging availability. Meanwhile, all CT scans are with contrast enhancement in portal venous phase. The in-plane pixel dimension of each CT scans varies from 0.7 to 1.2 mm. Each image is preprocessed by excluding outlier intensities beyond  $-1000$  and  $1000$  HU. Images slice thickness ranges from 2 to 4 mm. Each CT scans consists of 60 to 200 slices of  $512 \times 512$  pixels. For consistency consideration, we pre-processed all CT scans with a soft tissue window with a range of  $[-175, 250]$  HU, the window effect is studied in [24]. Intensities were normalized to  $[0, 1]$ .

**Pancreas Imaging Slices:** Clinically acquired CT scans usually have large variance in field of view. We implemented a pre-processing step with body part regression, which is a method to illustrate abdominal intrinsic structure [25]. The body part regression assists to automatically remove slices based on inconsistent volumes, and to localize abdominal

anatomies. We adopted the pre-trained model from the unsupervised regression network [25] to find slices in the pancreas region (scalar reference index score ranges from  $-1$  to  $1$  as indicated in [25]).

**Timeline of T2DM diagnosis and CT sessions:** To clarify the task of predicting future risk of T2DM, we obtained longitudinal CT sessions along with ICD-9 T2DM diagnosis code and date. We first assured the diagnosis date for each T2DM patient, then retrieved a CT session and EMR at least one year ahead of the ICD-9 date. A randomized controlled trial for a CT session and EMR was collected from a non-T2DM control group. The time interval between diagnosis date and CT session date ranges from 365 to 1690 days (mean: 456, median: 540). Only one CT session per patient is used in the study.

## 2.2 Abdominal Segmentations

To acquire volumetric measurements of pancreas tissue and obesity of patients, we perform abdominal segmentation on the pancreas [26, 27, 28], body wall/mask, and body fat from CT imaging [29]. In this paradigm, we use a dataset of 100 subjects from the MICCAI 2015 Multi-Atlas Labeling Challenge with 12 anatomies labels annotated by experts. In detail, we train a 3D multi-organ segmentation network [30] for segmenting the pancreas and inner/outer abdomen wall by multi-task learning. Each scan is down-sampled from  $[512, 512]$  to  $[168, 168]$  and normalized to consistent voxel resolution of  $[2 \times 2 \times 6]$ . The output and ground truth labels are compared using Dice loss [31]. We ignored the background loss in order to increase weights for anatomies. The final segmentation maps are up-sampled to original space with the nearest interpolation in order to spatially align with CT resolution. This approach is trained end-to-end, and the resulting segmentation is shown in Fig. 2. In order to extract fat volume measurements, we perform fuzzy c-means [32] on CT images for body mask and fat segmentation in an unsupervised scheme.

## 2.3 EMR Feature Extraction

A clinical EMR captured patient demographics, clinical notes, lab tests, medication codes, diagnosis codes, treatment procedures. For our cohort, we extracted the de-identified clinical phenotype base on a hierarchical categorization of ICD-9 (International Classification of Disease - 9) codes [33]. A pcode-groups a set of ICD-9 codes to imply a diagnosis category. For example, type II diabetes mellitus includes *250.00, 250.02, 250.40, 250.42, ..., etc.* We use an open-sourced PyPhewas tool to automatically generate the ICD-9 categories [17], given a list of clinical visits.

In this study, we investigate five different configurations of EMR features in predicting onset T2DM in an ablation study scheme:

**EMR1: Demographic.**—We extract demographic history for each de-identified patient involves age, sex, race, ethnicity, weight, height, and BMI. A vector of length 7 is formed for each subject. The risk factors are normalized from 0 to 1.

**EMR2: Pancreas volume + EMR1**—Pancreatic necrosis and acute pancreatitis mortality are major factors of insulin insufficiency. Herein, we calculate pancreas volume for each patient from the CT session for one year prior to diagnoses at predicting T2DM. The volumetric measurements are acquired from pancreas segmentation using 3D U-net as shown in Sect. 2.2.

**EMR3: Fat volume in L2 region + EMR2.**—Obesity increases the risk of systemic inflammatory response syndrome on T2DM prediction. Abdominal fat represented by waist circumference is better than single weight or BMI at characterizing metabolic syndrome [34]. We extract visceral and subcutaneous fat volume along the L2 section. Fat volumes are acquired from automatic fat/abdomen wall segmentation and distribution regressed by body part regression [25], where L2 slices are retrieved by body part regression scores in range  $-1$  to  $0$ .

**EMR4: Abdominal fat volume distribution + EMR3.**—To identify the utility of abdominal fat in prediction, we evaluated abdominal visceral fat volume and subcutaneous fat volume. Similar to EMR3, we extract abdominal slices by BPR scores in range of  $-6$  to  $5$ .

**EMR5: Glucose lab test + EMR4.**—Lab testing enables detection of prediabetes and suggestions for weight loss and other lifestyle changes to help delay or prevent T2DM. We retrieved the glucose test a year ahead of diagnoses associated with each patient into the EMR vector. In our study, 17 T2DM cases and 87 subjects from control are lack of the glucose test value, the missing value imputation is discussed in [35]. We present mean imputation that simply calculate the mean of the observed values for individuals who are nonmissing from T2DM cases and controls, respectively.

In order to project unimodal representation using the heterogeneity of multimodal data, we use a joint representation learning framework. In this work, we address the joint training problem by introducing an EMR-image multimodal network. As shown in Fig. 2, we include neural network for each modality that fused before the last fully connected and activation layer. Next, we fuse them at the last hidden layer, where the posteriors obtain the correlation between projected non-linear spaces of two modalities. Then this feature space is fixed and passed through a shallow linear layer to the targets. The networks are optimized by cross-entropy objective and stochastic gradient descent. The network comprises four 3D conv layers, followed by ReLU and pooling layers. The whole framework is trained in end-to-end manner and results in multimodal fusion directly.

### 3. Experiments

#### 3.1 Experimental Design

To perform each EMR configuration as well as bilinear multimodal network, we implemented experiments with cross validation on 997 subjects from the “case-control” study. To perform standard five-fold cross validation, we split dataset into five

complementary folds, each of which contains 200 cases (197 in last fold). For each fold evaluation, we use three folds as training, one fold as validation, and remaining as testing.

### 3.2 Implementation Details and Metric

The EMR-image multimodal network consists of four levels of convolution blocks, each block has a  $3 \times 3 \times 3$  convolutional layer, followed by rectified linear units (ReLU) and a max-pooling of  $2 \times 2 \times 2$  and strides of 2. The learning rate for the EMR-image multimodal network is set with  $1e-5$ . We use a batch size of 1 for all implementations and adopted the ADAM algorithm with SGD, momentum = 0.9. Implementations are performed using Pytorch 1.0 with NVIDIA Titan X GPU 12G memory and CUDA 9.0. The code of all experiments, including baseline methods, is implemented in Python 3.6 with Anaconda3. The area under the curve (AUC) is calculated from the receiver operating curve (ROC) for each experiment. ROC curves are generated by the true positive rate (TPR) against the false positive rate (FPR) at various threshold settings. AUC measures the entire two-dimensional area underneath the entire ROC curve.

### 3.3 Results and Analyses

The experimental result of all settings is shown in Table 1. We summarize the mean and standard deviation in accuracy, AUC, F1, recall, and precision for comparison. In the context of five EMR configurations, we can see that each risk factor contributes to improvements. In particular, improvements over demographics increases as risk factors get more diversity from images. We observe a higher improvement in adding body fat distribution. We also see consistent better performance using pancreas slices, showing the advantage of imaging biomarkers. In the multimodal analysis, we achieve the best accuracy at 86.19%, combining the pancreas slice and EMR5. Comparing with EMR5 or pancreas slices, our method achieved consistently better result, with 6.93% and 4.25% in AUC respectively. These imply the effectiveness of using heterogeneous data.

Figure 3 shows the ROC-AUC curves for all experiments. The green curve indicates the combination of EMR and pancreas slices in the training, we show that EMR-image multimodal framework for heterogeneous data analysis improves predictive power (AUC 0.8910) compare to brown curve (AUC 0.8485) and yellow curve (AUC 0.8217).

## 4. Discussion and Conclusion

In this work, we targeted at understanding of the association among T2DM, patient clinical history, tissue composition imaging as well as volumetric measurements. In this paper, we used a “case-control” study design for the onset of T2DM prediction task. We show that direct prediction using either EMR features or CT imaging enables prediction of risk factors for T2DM. We investigate five different configurations of EMR including demographics, pancreas volume measurements, visceral/subcutaneous fat volume, body fat distribution and glucose lab tests. Different factors contribute to improve AUC of the prediction. Meanwhile, we institute a holistic method for EMR features and CT images, and show improved performance over the base EMR features and imaging methods alone. The proposed method



builds upon the connections between different modalities (EMR and CT) through which we extrapolate a joint representation learning in the multimodal machine learning setting. The way we obtain the joint feature space represents a general means of exploitation based on multiple modalities. Specifically, jointly using imaging techniques, notes and quantitative measurements offers innovative understandings of the challenge. It also allows us to address problems systematically. We are excited to explore more mechanisms such as bilinear multimodal architecture [36] or dual attention [37], and in particular study the EMR-guided approaches in the future.

One of the limitations in this work focuses on the retrospective EMR study design. The ability of case-control design has been widely evaluated [21]. It sometimes is unable to detect very small relative risks from exposures and it tends to be more effective in large-scale, collaborative, multi-center studies. For example, compared to a previous study with AUC of 0.877 [8] using their institutional data, we achieved slightly higher AUC in our cohort (Table 1). The absolute performance on different institutional cohorts shows some relative risk factors could be typically regarded as low importance, but these factors may have higher importance in the population at large. Herein, the full cohort review is needed in the future. And to investigate subsequent studies to guarantee differences among controls do not confound results.

## Acknowledgements

This research is supported by Vanderbilt-12Sigma Research Grant, NSF CAREER 1452485, NIH 1R01EB017230 (Landman). This study was in part using the resources of the Advanced Computing Center for Research and Education (ACCRE) at Vanderbilt University, Nashville, TN. The imaging dataset(s) used for the analysis described were obtained from ImageVU, a research resource supported by the VICTR CTSA award (ULTR000445 from NCATS/NIH).

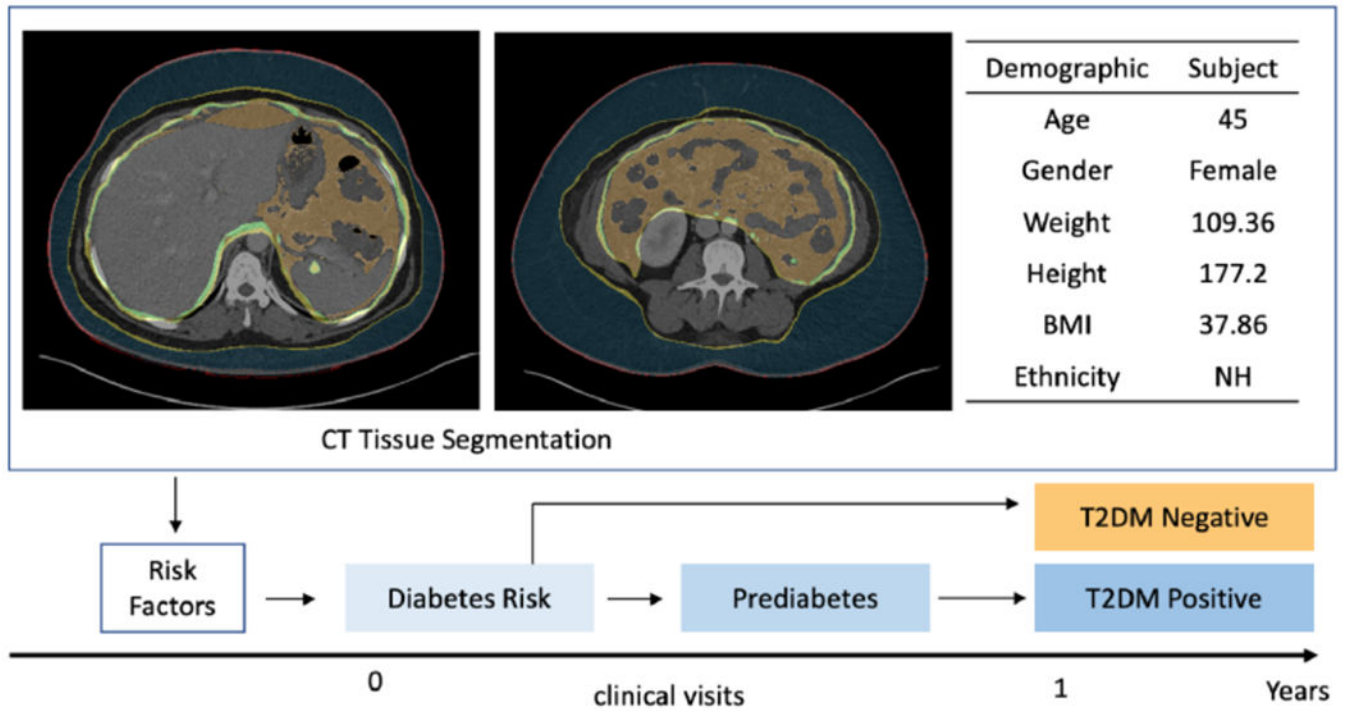
## References

1. Hales CN, Barker DJP: Type 2 (non-insulin-dependent) diabetes mellitus: the thrifty phenotype hypothesis. *Diabetologia* 35(7), 595–601 (1992). 10.1007/BF00400248 [PubMed: 1644236]
2. Chen L, Magliano DJ, Zimmet PZ: The worldwide epidemiology of type 2 diabetes mellitus—present and future perspectives. *Nat. Rev. Endocrinol* 8(4), 228–236 (2012)
3. Neeland IJ, et al.: Dysfunctional adiposity and the risk of prediabetes and type 2 diabetes in obese adults. *JAMA* 308(11), 1150–1159 (2012) [PubMed: 22990274]
4. Tognini G, Ferrozzì F, Bova D, Bini P, Zompatori M: Diabetes mellitus: CT findings of unusual complications related to the disease: a pictorial essay. *Clin. Imaging* 27(5), 325–329 (2003) [PubMed: 12932683]
5. Association AD: Diagnosis and classification of diabetes mellitus. *Diabetes Care* 37(Supplement 1), S81–S90 (2014) [PubMed: 24357215]
6. Fletcher B, Gulanick M, Lamendola C: Risk factors for type 2 diabetes mellitus. *J. Cardiovasc. Nurs.* 16(2), 17–23 (2002) [PubMed: 11800065]
7. Anderson AE, Kerr WT, Thames A, Li T, Xiao J, Cohen MS: Electronic health record phenotyping improves detection and screening of type 2 diabetes in the general United States population: a cross-sectional, unselected, retrospective study. *J. Biomed. Inform* 60, 162–168 (2016) [PubMed: 26707455]
8. Mani S, Chen Y, Elasy T, Clayton W, Denny J: Type 2 diabetes risk forecasting from EMR data using machine learning. In: *AMIA Annual Symposium Proceedings*, vol. 2012, p. 606. American Medical Informatics Association (2012) [PubMed: 23304333]

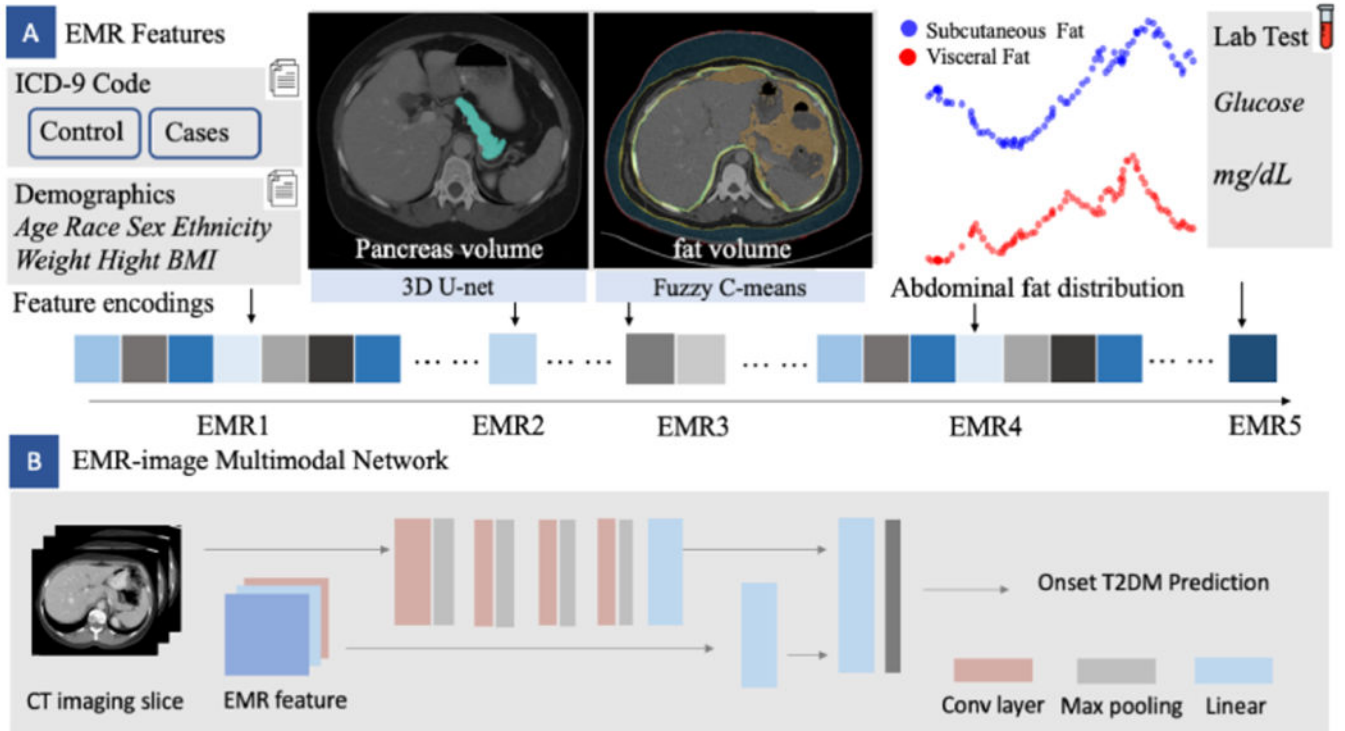
9. Virostko J, Hilmes M, Eitel K, Moore DJ, Powers AC: Use of the electronic medical record to assess pancreas size in type 1 diabetes. *PLoS ONE*, 11(7), e0158825 (2016) [PubMed: 27391588]
10. Kavakiotis I, et al.: Machine learning and data mining methods in diabetes research. *Comput. Struct. Biotechnol. J* 15, 104–116 (2017) [PubMed: 28138367]
11. Zheng T, et al.: A machine learning-based framework to identify type 2 diabetes through electronic health records. *Int. J. Med. Informatics* 97, 120–127(2017)
12. Garcia TS, Rech TH, Leitao CB: Pancreatic size and fat content in diabetes: a systematic review and meta-analysis of imaging studies. *PLoS ONE* 12(7), e0180911 (2017) [PubMed: 28742102]
13. Vu KN, Gilbert G, Chalut M, Chagnon M, Chartrand G, Tang A: MRI-determined liver proton density fat fraction, with MRS validation: comparison of regions of interest sampling methods in patients with type 2 diabetes. *J. Magn. Reson. Imaging* 43(5), 1090–1099 (2016) [PubMed: 26536609]
14. Zhang Z, Chen P, Shi X, Yang L: Text-guided neural network training for image recognition in natural scenes and medicine. *IEEE Trans. Pattern Anal. Mach. Intell* (2019)
15. Baltrušaitis T, Ahuja C, Morency L-P: Multimodal machine learning: a survey and taxonomy. *IEEE Trans. Pattern Anal. Mach. Intell* 41, 2 (2018)
16. Evans JA: Electronic medical records system. ed: Google Patents (1999)
17. Chaganti S, Bermudez C, Mawn LA, Lasko T, Landman BA: Contextual deep regression network for volume estimation in orbital CT. In: Shen D, Liu T, Peters TM, Staib LH, Essert C, Zhou S, Yap P-T, Khan A (eds.) *MICCAI 2019*. LNCS, vol. 11769, pp. 104–111. Springer, Cham (2019). 10.1007/978-3-030-32226-7\_12
18. Chaganti S, et al.: Electronic medical record context signatures improve diagnostic classification using medical image computing. *IEEE J. Biomed. Health Inform* 23(5), 2052–2062 (2018) [PubMed: 30602428]
19. Tang Y, et al.: Contrast phase classification with a generative adversarial network. arXiv preprint arXiv:1911.06395 (<http://arxiv.org.proxy.library.vanderbilt.edu/abs/1911.06395>) (2019)
20. Kulama E: Scanning protocols for multislice CT scanners. *Br. J. Radiol* 77(suppl\_1), S2–S9 (2004) [PubMed: 15546838]
21. Crombie IK: The limitations of case-control studies in the detection of environmental carcinogens. *J. Epidemiol. Community Health* 35(4), 281–287 (1981) [PubMed: 7338704]
22. Mann C: Observational research methods. *Research design II: cohort, cross sectional, and case-control studies*. *Emer. Med. J* 20(1), 54–60 (2003)
23. Landman B, Xu Z, Igelsias J, Styner M, Langerak T, Klein A: *MICCAI Multi-Atlas Labeling Beyond the Cranial Vault—Workshop and Challenge* (2015)
24. Huo Y, et al.: Stochastic tissue window normalization of deep learning on computed tomography. *J. Med. Imaging* 6(4), 044005 (2019)
25. Yan K, Lu L, Summers RM: Unsupervised body part regression via spatially self-ordering convolutional neural networks. In: *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pp. 1022–1025. IEEE (2018)
26. Xu Z, et al.: Efficient multi-atlas abdominal segmentation on clinically acquired CT with SIMPLE context learning. *Med. Image Anal* 24(1), 18–27 (2015) [PubMed: 26046403]
27. Xu Y, et al.: Outlier Guided Optimization of Abdominal Segmentation. arXiv2002.04098
28. Wang Y, Zhou Y, Shen W, Park S, Fishman EK, Yuille AL: Abdominal multi-organ segmentation with organ-attention networks and statistical fusion. *Med. Image Anal* 55, 88–102 (2019) [PubMed: 31035060]
29. Xu Z, Baucom RB, Abramson RG, Poulouse BK, Landman BA: Whole abdominal wall segmentation using augmented active shape models (AASM) with multi-atlas label fusion and level set,” in *Medical Imaging 2016: Image Processing*, 2016, vol. 9784: International Society for Optics and Photonics, p. 97840U
30. Çiçek Ö, Abdulkadir A, Lienkamp Soeren S., Brox T, Ronneberger O: 3D U-Net: learning dense volumetric segmentation from sparse annotation. In: *Ourselin S, Joskowicz L, Sabuncu Mert R.,*



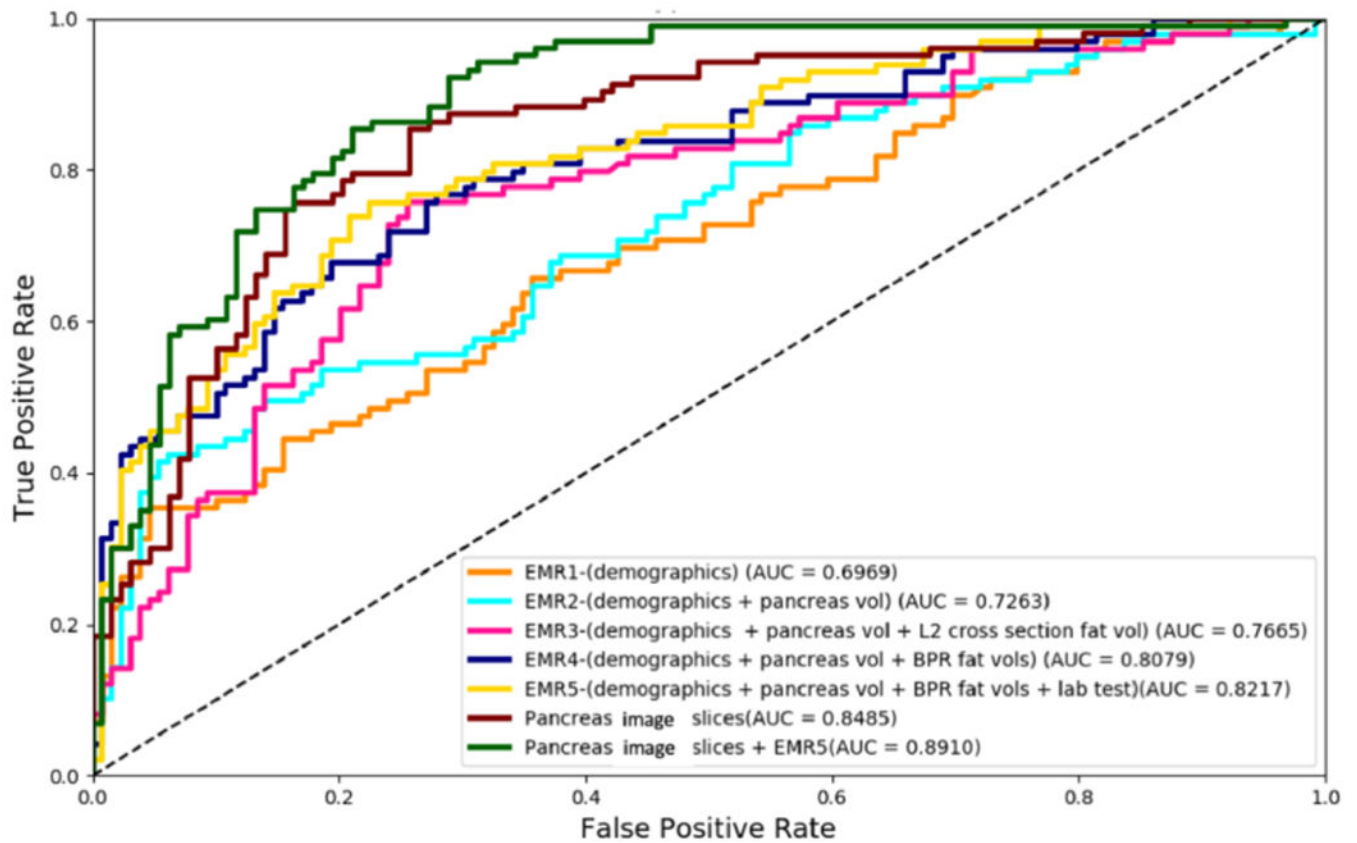
- Unal G, Wells W (eds.) MICCAI 2016. LNCS, vol. 9901, pp. 424–432. Springer, Cham (2016). 10.1007/978-3-319-46723-8\_49
31. Sudre Carole H., Li W, Vercauteren T, Ourselin S, Jorge Cardoso M: Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In: Jorge Cardoso M, et al. (eds.) DLMIA/MLCDS-2017. LNCS, vol. 10553, pp. 240–248. Springer, Cham (2017). 10.1007/978-3-319-67558-9\_28
  32. Bezdek JC, Ehrlich R, Full W: FCM: the fuzzy c-means clustering algorithm. *Comput. Geosci* 10(2–3), 191–203 (1984)
  33. Quan H, et al.: Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Med. Care* 43, 1130–1139 (2005) [PubMed: 16224307]
  34. Carey VJ, et al.: Body fat distribution and risk of non-insulin-dependent diabetes mellitus in women: the Nurses' Health Study. *Am. J. Epidemiol* 145, 7 (1997)
  35. Baraldi AN, Enders CK: An introduction to modern missing data analyses. *J. Sch. Psychol* 48(1), 5–37 (2010) [PubMed: 20006986]
  36. Mroueh Y, Marcheret E, Goel V: Deep multimodal learning for audiovisual speech recognition. In: 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 2130–2134. IEEE (2015)
  37. Zhang Z, Chen P, Sapkota M, Yang L: Tandemnet: Distilling knowledge from medical images using diagnostic reports as optional semantic references. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (2017)



**Fig. 1.** Example of CT images (subcutaneous fat in navy, visceral fat in brown). The flowchart shows from risk factors to T2DM diagnoses with multiple modalities (CT and demographic) at least a year ahead of diagnoses.



**Fig. 2.** Method framework. Part A indicates EMR features of five configurations. ICD-9 code is used as exclusion criteria for case-control design. The demographics is used as the base feature vector (EMR1). Pancreas and fat segmentation imply the volumetric measurements for EMR2 and EMR3, respectively. The abdominal fat distribution and glucose lab test are used to compose the EMR4 and EMR5 features. Part B shows our network architecture for combining EMR features and CT imaging slices for prediction of onset T2DM.



**Fig. 3.**

The receiver operating characteristic (ROC) curves of results on the prediction of onset T2DM. Area Under Curve (AUC) is shown in right bottom of figure. EMR features from 1 to 5 achieve barely satisfactory results, the use of pancreas slice and the proposed multimodal network performs better.

**Table 1.**

Result on the mean and standard deviation (std) (% , average (std) of five-fold cross validation). The *p-values* are calculated using McNemar test with respect to AUC in the row above each test.

Method	Accuracy	AUC	F <sub>1</sub>	Recall	Precision
EMR1	67.25(2.97)	69.69(2.45)	51.93(2.99)	39.28(4.64)	77.98(4.34)
EMR2	70.04(2.58)	72.63(2.13)	56.74(3.08)	49.72(4.04)	79.01(3.98)
EMR3	74.21(2.01)	76.65(1.75)*	59.82(2.76)	54.85(4.02)	80.06(3.34)
EMR4	76.93(1.96)	80.79(1.42)*	64.59(2.51)	59.81(3.54)	80.96(3.51)
EMR5	80.02(1.98)	82.17(1.31)*	67.42(2.43)	61.73(2.98)	82.79(3.64)
Pan slice	82.46(1.70)	84.85(1.19)	72.54(1.89)	64.49(2.74)	84.37(3.72)
Pan slice + EMR5	<b>86.19(1.45)</b>	<b>89.10(1.10)*</b>	<b>76.85(1.72)</b>	<b>67.61(1.98)</b>	<b>87.62(3.32)</b>

\* indicates significant with  $p < 0.01$ .

The best average results are shown in **bold**. Pan slice indicates pancreas slice.