

一般口演 | 医療データ解析

一般口演1

医療データ解析

2019年11月22日(金) 09:00 ～ 11:00 C会場 (国際会議場 2階国際会議室)

[2-C-1-02] BMIは糖尿病判別モデルに必須の情報か？

○藤田 怜一郎¹、奥村 健馬¹、山本 祐大¹、兵頭 勇己²、永田 桂太郎²、畠山 豊²、奥原 義保²（1. 高知大学医学部医学科先端医療学コース, 2. 高知大学医学部附属医学情報センター）

キーワード：Diabetes Mellitus, BMI, logistic regression model

【序論】2型糖尿病は進行によって予後不良となるため、早期発見による適切な治療開始が重要であり、日常診療で良く実施される検査を説明変数とする判別モデルの適用が考えられる。糖尿病の病態を表すために重要と考えられている BMIを用いるモデルが考えられるが、病院情報システム(HIS)では信頼できる形で構造化しておらず、データ収集のボトルネックとなっている。BMIを用いなくても判別精度が高いモデルを構築できれば、活用できるデータサイズは飛躍的に大きくなる。ここでは BMIが糖尿病判別モデルの判別能に与える影響について評価する。

【方法】高知大学附属病院の HISに1981年から2016年の間に保存された75gOGTTを行った入院患者のデータを対象とし、正常型(NM)、境界型(IGT)、糖尿病型(DM)に分類した。また判別モデルの作成に有意に影響を与えると考えられる検査項目を選択し、BMIは入院中看護データから求めた。NMと IGTおよび IGTと DMの関係についてロジスティック回帰モデルによる判別を行い、モデル間の AIC 及び AUCを比較した。

【結果】説明変数として年齢,RBC,Ht,Hb,Stab,TG,TP,Albを用いた判別モデルについて BMIを加えた場合の影響を評価したところ、NMと IGTでは男性の場合 AUCは0.752から0.766、AICは66.9から68.3に、女性の場合 AUCは0.802から0.804、AICは68.2から70.1と変化した。また説明変数を変更した場合及び IGTと DMの間の場合も同様の変化を示した。

【考察】AUCは微増にとどまり判別能向上への寄与はあまり見られず、AICもほぼ全てにおいて増加していることから、BMIによってモデルの改良がなされたとは考えづらい。関連する検査データが十分に存在している患者に対する糖尿病判別モデルには BMIが必ずしも必須ではないことが示唆された。

BMI は糖尿病判別モデルに必須の情報か

藤田 怜一郎^{*1}、奥村 健馬^{*1}、山本 祐大^{*1}
兵頭 勇己^{*2}、永田 桂太郎^{*2}、畠山 豊^{*2}、奥原 義保^{*2}

^{*1} 高知大学医学部医学科先端医療学コース、^{*2} 高知大学医学部附属医学情報センター

Is BMI the necessary factor for the discriminant model of Diabetes Mellitus?

Reiichiro Fujita^{*1}, Kemma Okumura^{*1}, Yudai Yamamoto^{*1}

Yuki Hyohdoh^{*2}, Keitaro Nagata^{*2}, Yutaka Hatakeyama^{*2}, Yoshiyasu Okuhara^{*2}

^{*1} Center for Innovative and Transitional Medicine, Kochi Medical School, Kochi University

^{*2} Center of Medical Information Science, Kochi Medical School, Kochi University

Abstract

In order to evaluate the effects of BMI on the discriminant model of Diabetes Mellitus, we targeted patients conducted OGTT and divided them into a normal type (NM), a boundary type (IGT) and Diabetes type (DM) by their OGTT results, and evaluated logistic regression models discriminating NM from IGT and IGT from DM that has lab test values as explanatory values. We compared AIC when the case which added BMI to each model as an explanatory variable or not. The result showed that AIC increased with almost all logistic models added BMI and suggested that BMI is not always the necessary factor for the discriminant model of Diabetes Mellitus.

Keywords: Diabetes Mellitus, BMI, logistic regression model

1. 結論

2 型糖尿病は働き盛りの世代で爆発的に増加しており、医療費の増加をはじめ社会に及ぼす影響も大きく、病態の進行によって予後不良となるため、早期発見による適切な治療開始が重要である。そのためには、患者にとって負担の少ない早期発見のための簡便な方法が望まれる。例えば、日常診療でよく実施される検査を説明変数とする判別モデルの適用が考えられる。そうしたモデルの構築のためには、多数の患者のデータが必要であり、病院情報システム(HIS)のような Real World Data の活用が期待される。しかしながら、ほとんどの HIS は糖尿病の病態を表すために重要と考えられている BMI を信頼できる形で構造化しておらず、データ収集のボトルネックとなっている。BMI を用いなくても判別精度が高いモデルを構築できれば、活用できるデータサイズは飛躍的に大きくなる。本研究では、BMI が糖尿病判別モデルの判別能に与える影響について評価する。

2. 方法

本研究では、高知大学医学部附属病院における病院情報システム(IMIS: Integrated Medical Information System)の匿名化データベース RYOMA2(Retrieval sYstem for Open Medical Analysis)に保存された 1981 年から 2016 年までの間に 75g 経口糖負荷試験(OGTT)を行った入院患者を対象としてデータの抽出を行なった。OGTT の検査結果より空腹時血糖値と負荷後 2 時間血糖値を用いて、糖尿病診断基準¹⁾に従い対象の患者を正常型(NM)、境界型(IGT)、糖尿病(DM)の 3 群に分類した。NM は空腹時値 110mg/dL 未満かつ 2 時間値 140mg/dL 未満を満たす患者群、DM は空腹時値 126mg/dL 以上或いは 2 時間値 200mg/dL 以上を満たす患者群、IGT は NM と DM いずれの基準も満たさない患者群とした。

また目的の判別モデルの作成に有意に影響を与えると考えられる検査項目のうち、スクリーニング検査として使用されるものとして、赤血球(RBC)、ヘモグロビン(Hb)、ヘマトクリット(Ht)、桿状核球(Stab)、中性脂肪(TG)、総蛋白(TP)、アルブミン(Alb)の 7 項目に加え、年齢および BMI(Body Mass Index)の計 9 項目を選択した。BMI は入院中看護データに保存された身長および体重の 2 値より $BMI(kg/m^2) = \text{体重}(kg) / \text{身長}^2(m^2)$ の定義に従い算出した。

NM と IGT および IGT と DM の判別を目的とするロジスティック回帰モデルを上記の 9 項目のうちから任意の項目を説明変数として選択することで組み合わせの異なるモデルをそれぞれ 5 種類作成した。その組み合わせは(BMI のみ)、(年齢、BMI)、(年齢、TP、TG、Alb、BMI)、(年齢、RBC、Ht、Hb、Stab、BMI)、(年齢、TP、TG、Alb、RBC、Ht、Hb、Stab、BMI)とした。ここでは、血球系の影響を考慮する目的として RBC、Ht、Hb、Stab の 4 項目を、生化学所見の影響を考慮する目的として TP、TG、Alb の 3 項目を選択し、これらの有無により説明変数の組み合わせを作成した。

BMI を説明変数として加えた場合の糖尿病判別モデルへの影響について、BMI を含むモデルと含まないモデルの AIC (赤池情報量規準)を比較することによりモデルの信頼性の評価を行い、AUC (Area under curve)を比較することにより判別モデルの判別精度評価を行った。なお、AUC の信頼区間は DeLong の方法で計算した。統計的有意水準は $p=0.05$ とした。これらの統計解析は R(ver.3.0.0)で行った。

3. 結果

3 群における男女ごとの年齢と BMI の分布を図 1 に示す。また、説明変数の組み合わせを変えてロジスティック回帰分析を行った結果を表 1 に示す。

年齢、RBC、Ht、Hb、Stab、TG、TP、Alb を用いた判別モデルについて BMI を説明変数として加えた場合と加えていない場合の影響を評価したところ、NMとIGTでは男性の場合 AUC は 0.752 から 0.766、AIC は 66.9 から 68.3 に、女性の場合 AUC は 0.802 から 0.804、AIC は 68.2 から 70.1 と変化した(表 1)。また IGTとDMでは男性の場合 AUC は 0.797 から 0.794、AIC は 55.0 から 57.0 に、女性の場合 AUC はいずれも 0.729、AIC は 68.2 から 70.1 と変化した(表 1)。説明変数を変更した場合も同様の変化を示した(表 1)。

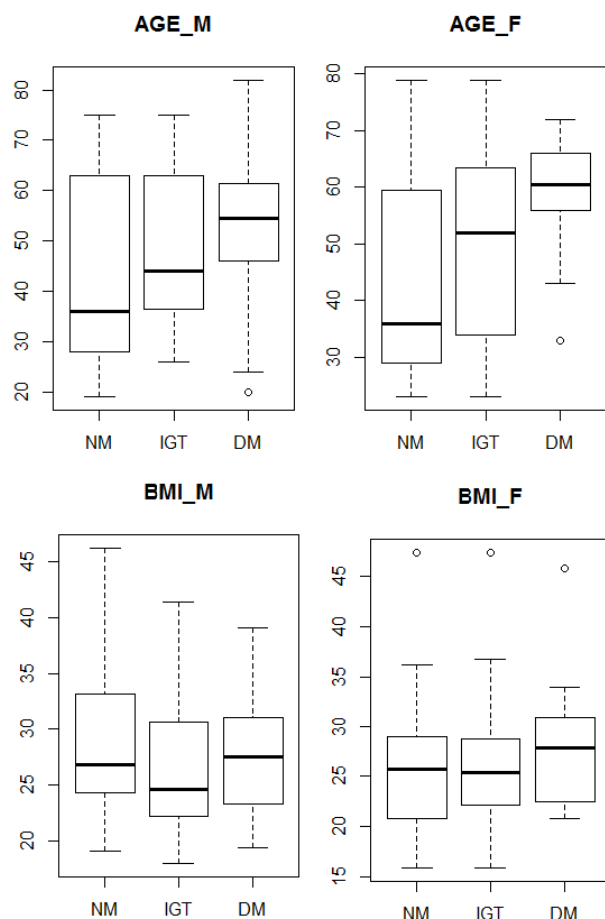


図1. 男女別、群別の年齢と BMI の分布

4. 考察

4.1 判別モデルの評価

今回作成した複数のモデルにおいて、BMI を説明変数として追加したことにより 1 つを除いた全てで AIC が微増したことから、判別モデル自体の精度の向上が認められなかった。AUC は増加しているがいずれのモデルにおいてもその程度は僅かであり、精度向上への寄与は小さいといえる。このことから少なくとも一時点の BMI は糖尿病の判別には必須ではないと考えられる。また、BMI を除く説明変数のみで AUC の一定の増加を認められるモデルも作成できたことから、説明変数の選択に機械学習等の解析手法を活用することを通じて精度の高い判別モデルの作成の可能性が見いだされたのではないかと考える。

また、他の説明変数を加えた際の AUC の変化にも注目する。RBC、Hb、Ht、Stab を加えた場合、TP、TG、Alb を加えた場合ではいずれか一方では顕著な AUC の増加は認めにくい

が、両方加えた場合は顕著な増加が認められる。男女別、判別の境界別において両者の影響の特徴に一定の差を認められなかった。このことは RBC、Hb、Ht、Stab と TP、TG、Alb の相互のバランスが糖尿病判別において何らかの有意な影響を与えていると考えられる。

4.2 年齢、BMI の分布

3 群における BMI 及び年齢の分布の違いを見ると、BMI は男女ともに顕著な差が認められなかったのに対し、年齢は重症な群ほど中央値が高くなっていることが分かる。このことは一時点の BMI では糖尿病の重症度の評価が難しいことを示唆しているといえ、また加齢にしたがってインスリン分泌能の低下及びインスリン抵抗性の増大が起きていると考えられる。

4.3 説明変数の採用

本研究で作成した判別モデルに説明変数として採用した検査項目の一つである Stab は、核の分葉・成熟が未分化な好中球であり、好中球の需要が増した時、すなわち免疫反応が誘発された時に急性に増加し得る。このため糖尿病患者にみられる高血糖状態が及ぼす血管傷害により Stab が増加し得るのではないかと考え採用した。

4.4 BMI の有用性

本研究で作成した糖尿病判別モデルにおいては、BMI が一時点の数値として扱われていたために有用な説明変数となり得なかったのではないかと考えられる。しかし、BMI は血液検査よりも簡易で負担が少ないこともあって、いつでも患者自身によっても計測可能であり、一目でおおよその肥満度を判断可能な点で有用な指標である。また、インスリンの分泌機能や感受性の変化は体重の増減として表れやすく、日々の生活習慣の改善やインスリン投与等による治療効果の予測など、長期的な糖尿病の病態の把握をするにあたっては、BMI が適していると考えられる。

4.5 本研究の限界

本研究の限界は大きく 2 つ挙げられる。1つは対象データが OGTT を実施した入院患者に限られることから、OGTT を実施するか否か、及び大学病院の入院患者か否かという点でバイアスがかかることである。OGTT 実施患者の BMI 分布を参照すればわかるように、男女ともいずれの群においても BMI の中央値は 25 程度かそれを超える値となっている。日本の BMI の平均値が男性で 23.7、女性は 22.5 であるという報告²⁾も考慮すると OGTT 実施患者は一般と比較して肥満傾向にあることが分かる。この BMI 分布の違いは OGTT の実施が糖尿病の病態の詳細な診断を目的としていることに起因しており、それらは肥満患者における糖尿病有病率の高さを根拠として実施されていると考えられる。

もう1つは、OGTT 実施患者のサンプル数があまり豊富ではなく条件の追加によりそれらを満たす患者の数が減少してしまうことである。この実施状況は OGTT 実施時の患者及び検査実施者の負担の大きさに由来すると考えられる。そのため、追加できる条件やその数に限界が生じ、有意な統計データを集めることが難しくなることが多くある。しかし、糖尿病の病態を、境界型を含む 3 群によって表すような検査としては OGTT が一番妥当な方法であり、詳細な糖尿病の病態の把握に重要である。

5. 結論

今回作成した糖尿病判別モデルに説明変数として BMI を加えても判別能の精度向上への明らかな寄与はみられなかった。一時点での糖尿病の判定には BMI は必ずしも必須ではないことが示唆された。BMI の指標としての扱い方について今後とも議論を続ける必要があるのではないだろうか。

参考文献

- 1) 日本糖尿病学会編, 糖尿病診療ガイドライン, 2016
- 2) 厚生労働省, 平成 28 年国民健康栄養調査報告, 2016

表1.NM 群と IGT 群及び IGT 群と DM 群の判別モデル

判別群	説明変数	性別	BMIの有無	AUC	AIC	感度	特異度	N
NM vs. IGT	BMI	男性	△	0.522(0.445-0.601)	300.64	97.2%	10.3%	214
		女性	△	0.576(0.528-0.623)	793.47	59.3%	58.9%	644
	AGE	男性	+	0.574(0.497-0.651)	298.79	60.7%	54.2%	214
			-	0.572(0.495-0.649)	297.15	83.2%	31.8%	
		女性	+	0.696(0.652-0.740)	736.25	57.3%	75.1%	644
			-	0.695(0.652-0.739)	744.41	70.9%	60.0%	
	AGE, TP,TG,Alb	男性	+	0.619(0.470-0.769)	89.89	40.9%	85.0%	62
			-	0.575(0.418-0.732)	89.06	59.1%	65.0%	
		女性	+	0.714(0.591-0.838)	96.91	51.7%	85.4%	70
			-	0.715(0.592-0.838)	94.91	51.7%	85.4%	
	AGE, RBC,Hb,Ht,Stab	男性	+	0.608(0.487-0.728)	127.37	38.6%	85.7%	86
			-	0.612(0.492-0.732)	126.27	54.5%	73.8%	
		女性	+	0.716(0.626-0.805)	172.76	81.6%	57.6%	134
			-	0.704(0.613-0.796)	171.98	79.6%	57.6%	
	AGE, RBC,Hb,Ht,Stab, TP,TG,Alb	男性	+	0.766(0.622-0.911)	68.29	81.2%	66.7%	43
			-	0.752(0.603-0.902)	66.87	56.2%	88.9%	
		女性	+	0.804(0.677-0.930)	70.15	95.0%	55.6%	47
			-	0.802(0.676-0.928)	68.20	75.0%	74.1%	
IGT vs. DM	BMI	男性	△	0.504(0.429-0.580)	320.42	82.0%	25.2%	229
		女性	△	0.523(0.458-0.588)	448.66	28.0%	80.4%	331
	AGE	男性	+	0.575(0.499-0.651)	315.50	90.2%	30.8%	229
			-	0.582(0.507-0.658)	314.15	92.6%	29.9%	
		女性	+	0.730(0.676-0.785)	393.65	85.6%	58.8%	331
			-	0.729(0.674-0.784)	395.10	81.8%	60.3%	
	AGE, TP,TG,Alb	男性	+	0.689(0.534-0.844)	71.02	48.0%	90.9%	47
			-	0.669(0.510-0.828)	71.00	48.0%	95.5%	
		女性	+	0.712(0.571-0.852)	79.71	60.0%	79.3%	54
			-	0.683(0.532-0.833)	78.82	64.0%	79.3%	
	AGE, RBC,Hb,Ht,Stab	男性	+	0.717(0.614-0.819)	128.83	51.0%	84.1%	95
			-	0.708(0.604-0.812)	128.26	60.8%	75.0%	
		女性	+	0.649(0.534-0.763)	127.89	82.1%	42.9%	88
			-	0.645(0.530-0.760)	125.93	59.0%	67.3%	
	AGE, RBC,Hb,Ht,Stab, TP,TG,Alb	男性	+	0.794(0.648-0.940)	57.03	80.0%	68.8%	36
			-	0.797(0.652-0.942)	55.04	80.0%	68.8%	
		女性	+	0.729(0.557-0.900)	59.37	85.7%	55.0%	34
			-	0.729(0.556-0.901)	57.58	71.4%	70.0%	

図2. 男女別及び BMI 有無別の年齢,RBC,Hb,Ht,Stab,TG,TP,Alb の ROC 曲線
 (左上:男性 BMI 有、右上:男性 BMI 無、左下:女性 BMI 有、右下:女性 BMI 無)

