表 6 JDCSグループ

主任研究者：山田信博(筑波大学)，**評価委員**：赤沼安夫(朝日生命糖尿病研究所)
分担協力研究者：衛藤雅昭，伊藤博史，網頭慶太(旭川医科大学)，赤沼安夫，菊池方利，野田光彦(朝日生命成人病研究所)，福本泰明，鷲見誠一(医療法人ガラシア病院)，清水靖久，小杉圭右(大阪警察病院)，星　充，渡會隆夫(大阪厚生年金病院)，竹村　芳，難波光義，宮川潤一郎，山崎義光(大阪大学)，阿部隆三(太田記念病院)，清野弘明(太田西ノ内病院)，石田俊彦(香川医科大学医学部)，藤田芳邦，矢島義忠(北里大学医学部)，名和田新(九州大学大学院医学研究院)，中埜幸治，中村直登(京都府立医科大学医学部)，岸川秀樹，豊永哲至(熊本大学)，野中共平，牧田善二，山田研太郎(久留米大学医学部)，武井　泉(慶応大学医学部)，貴田岡正史(公立昭和病院)，今泉昌利，東堂龍平(国立大阪病院)，山田研一(国立佐倉病院)，原納　優，吉政康直(国立循環器病センター)，野上哲史，西山敏彦(済生会熊本病院)，松岡健平(済生会糖尿病臨床研究センター)，梅津啓孝，仲野淳子(済生会福島総合病院)，片山茂裕(埼玉医科大学)，柏木厚典(滋賀医科大学)，古村幸雄(四国大学)，井上達秀(静岡県立総合病院)，石橋　俊(自治医科大学)，川上正舒(自治医科大学大宮医療センター)，河盛隆造(順天堂大学医学部)，大森安恵，河原玲子，佐藤麻子(東京女子医科大学)，北田俊雄，渡部良一郎(竹田綜合病院)，宮川高一(立川相互病院)，高橋和男，金塚　東，橋本尚武，齋藤　康(千葉大学医学部)，曽根博仁，山下亀次郎(筑波大学)，坂本美一，茂久田修(帝京大学市原病院)，田中　明(東京医科歯科大学)，佐々木敬(東京慈恵会医科大学)，門脇　孝，大須賀淳一，水野佐智子，藤井仁美，飯室　聡，大橋靖雄(東京大学)，藤田美明(東京都老人研究所)，井藤英喜(東京都多摩老人医療センター)，白井厚治(東邦大学附属佐倉病院)，高橋和眞(東北大学大学院医学系研究科)，村勢敏郎，小田原雅人(虎の門病院)，小林　正(富山医科薬科大学)，長瀧重信，赤澤昭一，川崎英二(長崎大学医学部附属病院)，堀出　饒，中村二郎(名古屋大学医学部)，及川眞一(日本医科大学)，林　洋一(日本大学医学部)，江草玄士，大久保政通，山根公則(広島大学医学部)，仲井継彦，笈田耕治(福井医科大学)，番度行弘(福井県済生会病院)，竹越忠美，若杉隆伸(福井県立病院)，豊岡重剛(福井赤十字病院)，小池隆夫(北海道大学医学部)，松島保久(松戸市立病院)，布目英男(水戸済生会総合病院)，豊島博行(箕面市立病院)，高橋秀夫(みなみ赤塚クリニック)，川崎　良，山下英俊(山形大学)，関原久彦(横浜市立大学医学部)，西川哲男(横浜労災病院)，南條輝志男(和歌山県立医科大学)
※当時の所属を含む．

ため，JDCS でメタボリックシンドローム診断基準ならびにその構成因子を満たした患者の，満たさなかった患者に対する心血管疾患リスクを算出した[25,39](表 5)．

その結果，女性糖尿病患者では NCEP 基準によりメタボリックシンドロームと診断された患者とそうでない患者との間では心血管疾患リスクの有意な違いはみられず，男性糖尿病患者では WHO 基準によりメタボリックシンドロームと診断された患者とそうでない患者との間に，心血管疾患リスクの有意な違いはみられなかった[25]．さらに，男性ではメタボリックシンドロームの診断基準を満たすより，むしろその一構成因子である“トリグリセリド上昇(150 mg/dl 以上)”のみを満たした場合のリスク上昇度のほうが大きいことが示唆された[25]．高トリグリセリド血症は香港人糖尿病患者においても独立した心血管リスクファクターであり[40]，東アジア人糖尿病患者の心血管リスクファクターとしての重要性が示唆される．

その後発表された日本と国際糖尿病連盟[41]の診断基準は，いずれも腹囲がメタボリックシンドローム診断の必須項目となっている．すなわち，腹囲が男性 85 cm，女性 90 cm(日本人の場合)を超えなければメタボリックシンドロームとは診断されない．両診断基準はトリグリセリドと HDL-コレステロールの扱い以外はほぼ同一であるが，IDF 診断基準を JDCS 患者にあてはめたところ，従来の WHO，NCEP 診断基準よりむしろ心血管疾患予知能が低く，あてはまった患者の虚血性心疾患・脳卒中ハザード比は有意に上昇していなかった[39]．

本研究の臨床的・社会的意義

本研究の意義のひとつは，前述のように非欧米人の糖尿病患者を対象にしたはじめての大規模臨床介入研究であるという点である．また，生活習慣改善を主体とした比較的緩やかな介入がわずかとはいえ，長期の HbA_{1C} 値低下を生み出すことも示した．また，欧米人患者とは異なる日本人患者のさまざまな病態的特徴が見出され，日本人患者によるエビデンスの必要性が示されたことも重要

な意義と思われる.

研究のインフラやマネジメントで苦労・工夫した点

JDCS 開始時には,わが国ではまだ無作為割付によるメガスタディが一般的でなく,国民皆保険が整備された社会背景や国民性からも欧米的手法はなじまないことが予想された.さらに,Kumamoto Study(「サイドメモ 3」参照)や DCCT などの先行研究で,血糖コントロールが細小血管合併症発症に大きな影響をもつことが知られていたため,従来の研究のように薬物を用いて強化治療群と対照群の治療内容に大差をつけることは許されない状況であり,生活習慣修飾をメインとした強化介入が採用された.

本研究のように長期にわたる大規模介入試験では主治医や患者の移動に伴う登録症例の脱落が起こりやすく,その防止にはとくに努力を要した.長期にわたって主治医・スタッフの間でモチベーションが維持されるように,定期的なニュースレターの発行なども行っている.

おわりに

JDCS は日本の糖尿病専門医が協力してつくり上げた貴重なデータベースであり,これを最大限に活用して将来の日本の糖尿病診療に役立つエビデンスを生み出していかなければならない.欧米人とは基礎的病態がかなり異なる日本人の糖尿病の診療や保健施策においては日本人患者のデータから得たエビデンスが求められており,その意味でも JDCS の今後の展開が期待される.

謝辞:全国の糖尿病専門施設の協力を得て実施中の厚生労働科学研究 JDCS にご参加いただいている多くの先生方・関係者・患者のご尽力に深謝いたします.本研究の統計解析は,東京大学大学院医学部生物統計学教室の大橋靖雄教授,田中佐智子先生(現国立がんセンター)により行われています.

サイドメモ 3

Kumamoto Study

2 型糖尿病患者において血糖コントロールが網膜症や腎症などの血管合併症発症を抑制しうることを,小規模ながら世界に先がけて示した日本発の重要な臨床研究である.2 型糖尿病 110 人を化療法(頻回インスリン注射)と従来治療(インスリン 1~2 回注射)とに分けて比較した.

文献

1) Roglic, G. et al.: *Diabetes Care*, **28**: 2130-2135, 2005.
2) Wild, S. et al.: *Diabetes Care*, **27**: 1047-1053, 2004.
3) Karter, A. J.: *Diabetes Care*, **26**: 2189-2193, 2003.
4) Abate, N. and Chandalia, M.: *J. Diabetes Complications*, **17**: 39-58, 2003.
5) Lanting, L. C. et al.: *Diabetes Care*, **28**: 2280-2288, 2005.
6) Lee, E. T. et al.: *Diabetologia*, **44**(Suppl. 2): S3-13, 2001.
7) Pan, X. R. et al.: *Diabetes Care*, **20**: 537-544, 1997.
8) Eriksson, J. et al.: *Diabetologia*, **42**: 793-801, 1999.
9) Diabetes Prevention Program Research Group: *N. Engl. J. Med.*, **346**: 393-403, 2002.
10) Norris, S. L. et al.: *Diabetes Care*, **24**: 561-587, 2001.
11) Sone, H. et al.; JD Study Group: *Horm. Metab. Res.*, **34**: 509-515, 2002.
12) Sone, H. et al.; Japan Diabetes Complication Study Group: *Horm. Metab. Res.*, **35**: 334-335, 2003.
13) Norris, S. L. et al.: *Diabetes Care*, **25**: 1159-1171, 2002.
14) Sone, H. et al.: *Diabetes Care*, **25**: 2115-2116, 2002.
15) Nelson, K. M. et al.; NHANES Ⅲ: Diet and exercise among adults with type 2 diabetes: *Diabetes Care*, **25**: 1722-1728, 2002.
16) Sone, H. et al.; Japan Diabetes Complication Study Group: *Lancet*, **361**: 85, 2003.
17) Saydah, S. H. et al.: *JAMA*, **291**: 335-342, 2004.
18) Tesfaye, S.; EURODIAB Prospective Complications Study Group: *N. Engl. J. Med.*, **352**: 341-350, 2005.
19) Sone, H. et al.: *N. Engl. J. Med.*, **352**: 1925-1927, 2005.
20) Kubo, M. et al.: *Stroke*, **34**: 2349-2354, 2003.
21) UK Prospective Diabetes Study(UKPDS) Group: *Lancet*, **352**: 837-853, 1998.
22) Howard, A. A. et al.: *Ann. Intern. Med.*, **140**: 211-219, 2004.
23) Sone, H. et al.: *Ann. Intern. Med.*, **141**: 408-409, 2004.

24) Mitka, M. : *JAMA*, **293** : 1845-1846, 2005.
25) Sone, H. et al. : *Diabetes Care*, **28** : 1463-1471, 2005.
26) Kemp, T. M. et al. : *Diabetes Care*, **28** : 1490-1492, 2005.
27) Comaschi, M. et al. ; SFIDA Study Group—Italian Association of Diabetologists (AMD) : *Nutr. Metab. Cardiovasc. Dis.*, **15** : 204-211, 2005.
28) Sone, H. et al. : *Am. J. Med.*, **117** : 711-712, 2004.
29) Howard, B. V. et al. : *Diabetes Care*, **29** : 391-397, 2006.
30) Alexander, C. M. et al. ; Third National Health and Nutrition Examination Survey (NHANESⅢ) ; National Cholesterol Education Program (NCEP) : *Diabetes*, **52** : 1210-1214, 2003.
31) Bruno, G. et al. : *Diabetes Care*, **27** : 2689-2694, 2004.
32) Hillier, T. A. et al. : *Diabetes Care*, **28** : 2258-2260, 2005.
33) Guzder, R. N. et al. : *Diabetologia*, **49** : 49-55, 2006.
34) Ko, G. T. et al. : *Diabetes Obes Metab.*, **8** : 94-104, 2006.
35) Juutilainen, A. et al. : *Diabetologia*, **49** : 56-65, 2006.
36) Isomaa, B. et al. : *Diabetes Care*, **24** : 683-689, 2001.
37) Saely, C. H. et al. : *J. Clin. Endocrinol. Metab.*, **90** : 5698-5703, 2005.
38) Kahn, R. et al. ; American Diabetes Association ; European Association for the Study of Diabetes : *Diabetes Care*, **28** : 2289-2304, 2005.
39) Sone, H. et al. ; Japan Diabetes Complications Study (JDCS) Group : *Diabetes Care*, **29** : 145-147, 2006.
40) Chan, W. B. et al. : *Diabetes Metab. Res. Rev.*, **21** : 183-188, 2005.
41) Alberti, K. G. et al. ; IDF Epidemiology Task Force Consensus Group : *Lancet*, **366** : 1059-1062, 2005.

＊＊次号の特集予告(220巻4号)＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊

◆眠りの科学——動物モデルによる睡眠覚醒研究

(企画：粂　和彦／熊本大学発生医学研究センター幹細胞制御分野)

＊＊＊

睡眠は，科学的には未解明の部分ばかりである．睡眠は意識を失うという現象を伴うが，この意識(心)がどのように脳でつくられるかがよくわかっていないのも大きな理由である．また，睡眠はマクロレベルで脳活動を調べる脳波により定義され，ある規模のネットワーク単位で起こる現象であり，ひとつの神経細胞レベルでは，“眠る”という現象は説明できない．ミクロレベルで個々のシナプス伝達の仕組みやニューロン活動に関しての研究は進んでも，マクロとミクロの研究にまだ断絶があるということである．ノンレム睡眠とレム睡眠により構築される睡眠は哺乳類・鳥類と一部の爬虫類にしか存在せず，両生類・魚類には原始的睡眠しかない．ところが，昆虫などの無脊椎動物にさえも睡眠類似行動が認められることが示され，意外なことにヒトの睡眠と似た部分もたくさんみつかり，これらの動物モデルを使った研究が注目されている．睡眠研究は，日本が世界をリードしている分野のひとつである．本特集では臨床医学からはすこし離れ，動物モデルを使ってユニークな研究を展開している研究者の方がたに，睡眠研究の面白さと最先端の知見をご紹介いただく．

Japan Diabetes Complications Study (JDCS)
－日本人２型糖尿病患者の大規模臨床研究－

[1]お茶の水女子大学　人間文化創成科学研究院
ライフサイエンス専攻　生活習慣病医科学
助教授　**曽根　博仁**[1,2]　Hirohito Sone

[2]筑波大学医学専門学群　内分泌代謝・糖尿病内科
教授　**山田　信博**　Nobuhiro Yamada

朝日生命成人病研究所
名誉所長　**赤沼　安夫**　Yasuo Akanuma

Japan Diabetes Complications Study Group（後記）
JDCSグループ

わが国の２型糖尿病患者は，生活習慣の欧米化と人口の高齢化を背景に，戦後一貫して急増してきた。糖尿病は，特異的な三大合併症（網膜症，腎症，神経障害）をもたらすだけでなく，冠動脈疾患，脳卒中などの動脈硬化疾患（すなわち大血管合併症）のリスクも数倍以上高める。２型糖尿病とその血管障害は世界規模でみても深刻な健康問題で，全世界死亡の5.2％にあたる290万人の過剰死亡の原因と見積もられている[1]。

２型糖尿病の病態には，人種・民族間で違いがあることが知られており，例えば，東アジア人糖尿病患者は欧米人患者より冠動脈疾患発症率がかなり低い[2,3]。そのため可能であれば，日本人患者の診療には，日本人患者データに基づくエビデンスを用いる方がよいと思われる。さらに日本人患者のエビデンスは，世界の糖尿病患者の約４割を占めるアジア人患者[4]の診療にも役立つ可能性が高い。しかし，糖尿病の大規模臨床研究の多くは欧米白人患者を対象にしたもので，アジア人患者を多く含んだ研究は少なかった。

1996年に開始されたJapan Diabetes Complications Study（JDCS）[5]は，日本人２型糖尿病患者を対象にした大規模臨床介入研究で，現在も継続中である。デザインとしては，無作為割り付け比較対照介入試験で，主な介入内容が生活習慣介入であるためオープンラベル法となっている。登録患者は，それまでの外来治療を継続する「非介入群」と，主治医と協

表１．JDCS（登録時）と英国，米国の糖尿病患者の臨床的特徴の比較（文献８，９より改変引用）

	JDCS（日本）	UKPDS（英国）	NHANES（米国）
	2,205	2,015	441
年齢（歳）	59	62	59
糖尿病罹患期間（年）	11	9	13
血圧（mmHg）	132/77	140/80	135/72
空腹時血糖値（mg/dL）	158	147	データなし
HbA_{1c}（%）	7.7	7.9	7.8
総コレステロール（mg/dL）	201	205	209
トリグリセリド（mg/dL）	125	137	データなし
Body mass index（BMI）(kg/m^2)	23.1	29.4	32.3
一般人口の平均BMI（kg/m^2）	22.7	24.1	28.5

力して外来や電話における生活指導強化により，生活習慣改善を目指す「介入群」の2群に割り付けられている。介入内容が現行ガイドラインに沿った緩やかなもので，治療が群間で極端には違わないため，全体を一つの集団とみなして全体的な特徴も検討されている。

対象患者は，全国の糖尿病専門施設59ヵ所に通院するHbA1c 6.5％以上の2型糖尿病患者2,205名である（前増殖性以上の網膜症や心血管疾患など進行した合併症をもつ患者は除外された）。臨床的特徴を表1に示したが，都市部病院に通院する発症後10年くらいの典型的な2型糖尿病患者像と思われる。これらの患者の血糖・血圧・血清脂質など多くの項目について，年1回の調査が続けられ，特に合併症については，あらかじめ定められた診断基準に基づき，専門委員により発症の判定が行われている。

これまでのJDCS研究

現在までに，3年目の中間結果[6]が発表されている。今後，各合併症の発症率やリスクファクターに関する本格的な解析結果が続く予定であるが，これまでになされた予備的な探索解析からも，欧米人糖尿病患者とは異なる日本人患者の様々な特徴が垣間見える。

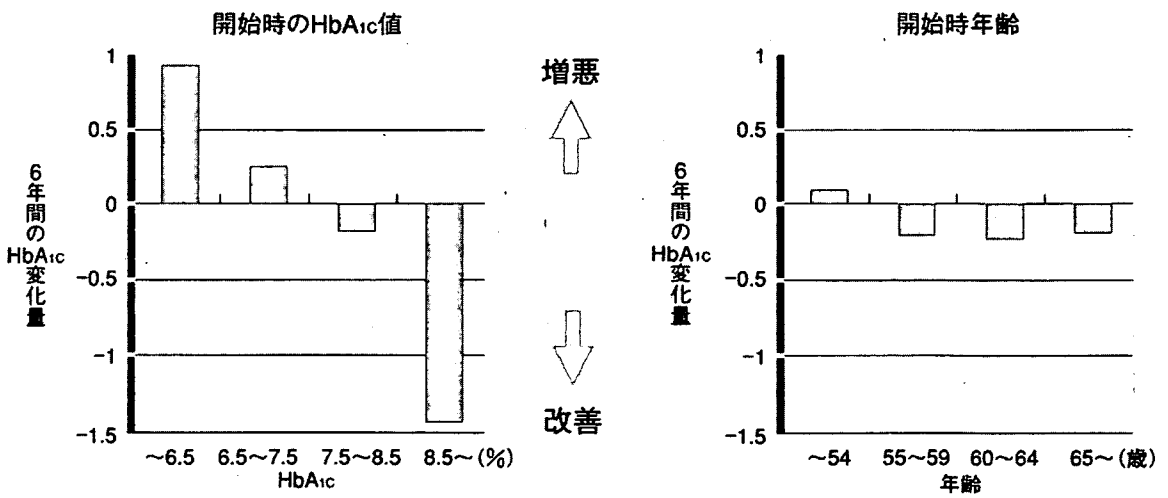
1. 肥満

欧米では，糖尿病患者の大部分は肥満があるとされる[7]。糖尿病罹患期間，年齢，血糖コントロールなどが類似するJDCS登録患者と，英米の糖尿病患者とを比較したところ[8,9]，確かに英米人糖尿病患者は著明な肥満を認め，しかも糖尿病患者の平均BMIは一般人口平均よりかなり高かった。しかし，JDCS登録日本人患者の平均的BMIは正常範囲であり，なおかつ一般人口のBMIとほとんど同等であった（表1）。日本人糖尿病患者では，欧米人患者よりインスリン抵抗性が著明でないかわりに，インスリン分泌能は低下しやすいことが従来から指摘されてきたが，そのことを裏付ける結果といえる。

2. 血糖のコントロール

JDCS登録者全体を，登録時のHbA1c値または年齢で層別化し，その後6年間のHbA1c変化量を検討したところ，登録時HbA1c値7.5％以上の患者または，登録時年齢55歳以上では，いずれもその後6年間でHbA1cの改善がみられたのに対して，登録時HbA1c値7.5％未満の患者または，登録時年齢54歳以下では，いずれも，その後6年間でHbA1cが増悪

図1. JDCS開始時のHbA1c値および年齢別にみたその後6年間のHbA1c変化（文献5より改変引用）



した(図1)。このことは，比較的若年または初期の患者で長期的な血糖コントロール増悪がみられやすかったことを示唆している。大血管障害(心血管疾患)の発症リスクが，糖尿病初期から上昇していることは日本人患者でも示されており[10]，若年初期の患者に対する指導を強化する必要性が示唆される。

3. 血清脂質のコントロール

糖尿病合併症のうちでも，特に大血管障害の抑制には血糖のみならず，血清脂質や血圧のコントロールが肝要である。LDLコレステロール100 mg/dL未満の患者に対する160 mg/dL以上の患者の冠動脈疾患リスクは3.1倍(95％信頼区間：1.6～6.3)になり，収縮期血圧130 mmHg未満の患者に対する150 mmHg以上の患者の脳卒中リスクは2.2倍(95％信頼区間：1.2～3.9)であった。しかし，収縮期血圧130 mmHg未満，かつ拡張期血圧80 mmHgを満たすコントロール良好者は，糖尿病患者の4割に達せず，総コレステロール200 mg/dL以下のコントロール良好者も半数に満たなかった。

図2. JDCSにおける開始4年後までの糖尿病網膜症発症
(開始時のHbA1cのレベルによって層別化した場合)

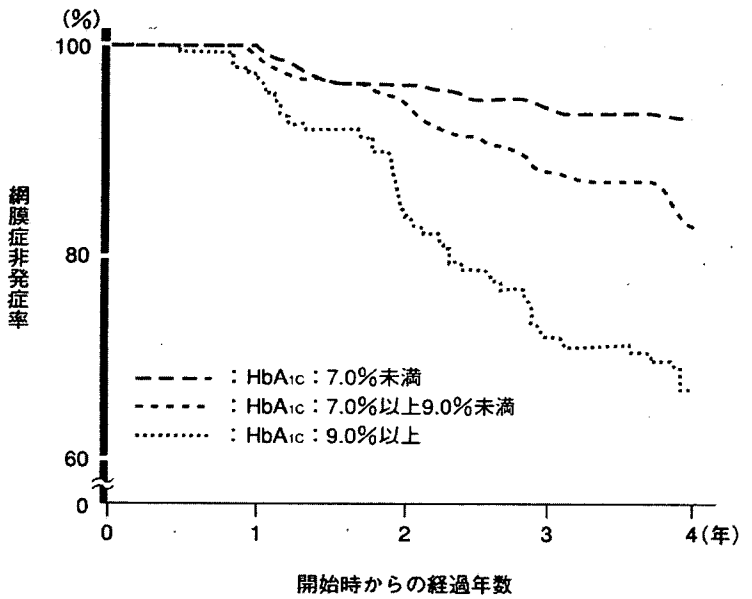


表2. 日本および欧米の糖尿病患者におけるアルコール摂取と心血管合併症の関係(文献16より)

	Howard, et al.(文献15より)	JDCS(文献16より改変)	
	冠動脈疾患	冠動脈疾患	脳卒中
飲酒なし	100％	100％	100％
38gエタノール(＝日本酒1.5合)までの飲酒	45～66％	125％	117％
それ以上の飲酒	143％	119％	198％*

(*飲酒なしに対してp＜0.05)

4. 糖尿病網膜症

網膜症発症率は，開始時HbA1c値の違いにより明らかな差が認められ，血糖コントロールの重要性が確認された(図2)。

5. 糖尿病腎症

網膜症と同様に血糖コントロールの影響は強く，開始時HbA1c 7％未満の層と比較すると，HbA1c 9％以上の層のリスクは，網膜症で5.2倍，腎症で4.5倍にも達することが示された。血糖のコントロールも重要であることが確かめられている。

6. 糖尿病神経障害

最近のヨーロッパの1型糖尿病患者を対象にした研究[11]では，HbA1cと罹病期間のほかに，これらで調整した解析では，多くの動脈硬化リスクファクターが神経障害のリスクファクターでもあることが判明した。しかし，JDCS登録患者で同様に検討してみると，大部分の動脈硬化リスクファクターは神経障害のリスクファクターにはなっておらず，神経障害のリスクファクターは，糖尿病の型や人種により大きく異なることが示された[12]。

7. 大血管障害

冠動脈疾患の患者1,000人あたりの年間発症率は8.8(男性10.6，女性6.8)で，久山町研究[13]の一般住民のデータ(男性3.48，女性1.81)と比較してかなり高値を示した。JDCSでは，冠動脈疾患の年間発症率の方が，脳卒中の患者1,000人あたりの年間発症率は7.9(男性8.5，女性7.0)よりむしろ高かった。英国のUKPDS[14]と比較すると，JDCSでは虚血性心疾患は約半数，脳血管障害はJDCSがやや多かった。

また，欧米人糖尿病患者のメタアナリシス[15]では，適度(純エタノール換算で1日38g以下)のアルコール摂取は，冠動脈疾患の抑制を有することが示された。しかし，JDCS登録患者ではそのような効果は認められず[16](表2)，日本人2型糖尿病患者に対しては，例え適量でも飲酒を積極的にすすめる根拠は薄いと思われた。

8. メタボリックシンドロームの意義

メタボリックシンドロームは，インスリン抵抗性を基盤とした心血管リスクファクターの重積が，心血管疾患を相乗的に増加させる病態である。欧米の前向き研究では，メタボリックシンドローム合併糖尿病患者は，非合併糖尿病患者より心血管疾患リスクが高いことが示されている。さらに，メタボリックシンドロームは非糖尿病者においても心血管イベントのリスクを高めるが，2型糖尿病患者の方がその影響が強い(リスク上昇度が大きい)ことも報告されている。

そこで，日本人2型糖尿病患者において，メタボリックシンドロームが心血管疾患の発症予知に有用であるかを検討するため，JDCSにおいて各種のメタボリックシンドローム診断基準を満たした患者の，満たさなかった患者に対する心血管疾患リスクを算出したところ，欧米人患者と異なり日本人患者では，メタボリックシンドローム診断が必ずしも心血管疾患予測に最適とはいえないことが示された[17, 18]。日本人糖尿病患者の心血管疾患リスク評価に適した指標の開発が望まれる。

JDCSの意義

本研究の意義の一つは，非欧米人の糖尿病患者を対象にした初めての大規模臨床介入研究であるという点である。その中で，欧米人患者とは異なる日本人患者の様々な病態的特徴が見いだされ，日本人糖尿病患者の診療には，日本人患者のエビデンスが必要であることを強く示唆する結果が得られている。JDCSは，日本の多くの糖尿病専門医・患者・関係者が協力して作り上げた貴重なデータベースであり，これを最大限に活用して，将来の日本の糖尿病診療に役立つエビデンスを今後さらに生み出すことが期待されている。

謝　辞

本研究の統計解析は，東京大学大学院生物統計学/疫学教室の大橋靖雄教授・田中佐智子先生（現：国立がんセンターがん予防・検診研究センター情報研究部）が担当されています。また本研究は，下記糖尿病専門施設の共同研究であり，ご参加いただいている多くの先生方・関係者・患者さんのご尽力に深謝いたします。

JDCS グループ

主任研究者：山田信博（筑波大学）

評価委員：赤沼安夫（朝日生命糖尿病研究所）

分担協力研究者：衛藤雅昭，伊藤博史，網頭慶太（旭川医科大学），赤沼安夫，菊池方利，野田光彦（朝日生命成人病研究所），福本泰明，鷲見誠一（医療法人ガラシア病院），清水靖久，小杉圭右（大阪警察病院），星　充，渡會隆夫（大阪厚生年金病院），竹村　芳，難波光義，宮川潤一郎，山崎義光（大阪大学），阿部隆三（太田記念病院），清野弘明（太田西ノ内病院），石田俊彦（香川医科大学医学部），藤田芳邦，矢島義忠（北里大学医学部），名和田新（九州大学大学院医学研究院），中埜幸治，中村直登（京都府立医科大学医学部），岸川秀樹，豊永哲至（熊本大学）野中共平，牧田善二，山田研太郎（久留米大学医学部），武井泉（慶応大学医学部），貴田岡正史（公立昭和病院），今泉昌利，東堂龍平（国立大阪病院），山田研一（国立佐倉病院），原納　優，吉政康直（国立循環器病センター），野上哲史，西山敏彦（済生会熊本病院），松岡健平（済生会糖尿病臨床研究センター），梅津啓孝，仲野淳子（済生会福島総合病院），片山茂裕（埼玉医科大学），柏木厚典（滋賀医科大学），吉村幸雄（四国大学），井上達秀（静岡県立総合病院），石橋　俊（自治医科大学），川上正舒（自治医科大学大宮医療センター），河盛隆造（順天堂大学医学部），大森安恵，河原玲子，佐藤麻子（東京女子医科大学），北田俊雄，渡部良一郎（竹田綜合病院），宮川高一（立川相互病院），高橋和男，金塚　東，橋本尚武，齋藤　康（千葉大学医学部），曽根博仁，山下亀次郎（筑波大学），坂本美一，茂久田修（帝京大学市原病院），田中　明（東京医科歯科大学），佐々木敬（東京慈恵会医科大学），門脇　孝，大須賀淳一，水野佐智子，藤井仁美，飯室　聡，大橋靖雄（東京大学），藤田美明（東京都老人研究所），井藤英喜（東京都多摩老人医療センター），白井厚治（東邦大学附属佐倉病院），高橋和眞（東北大学大学院医学系研究科），村勢敏郎，小田原雅人（虎の門病院），小林　正（富山医科薬科大学），長瀧重信，赤澤昭一，川崎英二（長崎大学医学部附属病院），堀田　饒，中村二郎（名古屋大学医学部），及川眞一（日本医科大学），林　洋一（日本大学医学部），江草玄士，大久保政通，山根公則（広島大学医学部），仲井継彦，笈田耕治（福井医科大学），番度行弘（福井県済生会病院），竹越忠美，若杉隆伸（福井県立病院），豊岡重剛（福井赤十字病院），小池隆夫（北海道大学医学部），松島保久（松戸市立病院），布目英男（水戸済生会総合病院），豊島博行（箕面市立病院），高橋秀夫（みなみ赤塚クリニック），川崎　良，山下英俊（山形大学），関原久彦（横浜市立大学医学部），西川哲男（横浜労災病院），南條輝志男（和歌山県立医科大学）（当時の所属を含む）

■文献

1) **Roglic G, et al.**：Diabetes Care 2005；**28**：2130－2135.
2) **Karter AJ**：Diabetes Care 2003；**26**：2189－2193.
3) **Abate N, Chandalia M**：J Diabetes Complications 2003；**17**：39－58.
4) **Wild S, et al.**：Diabetes Care 2004；**27**：1047－1053.
5) **Sone H, et al.**：Horm Metab Res 2002；**34**：509－515.
6) **Sone H, et al.**：Horm Metab Res 2003；**35**：334－335.
7) **Nelson KM, et al.**：Diabetes Care 2002；**25**：1722－1728.
8) **Sone H, et al.**：Lancet 2003；**361**：85.
9) **Saydah SH, et al.**：JAMA 2004；**291**：335－342.
10) **Fujishima M, et al.**：Diabetes 1996；**45** (Suppl 3)：S14－16.
11) **Tesfaye S, et al.**：N Engl J Med 2005；**352**：341－350.
12) **Sone H, et al.**：N Engl J Med 2005；**352**：1925－1927.
13) **Kubo M, et al.**：Stroke 2003；**34**：2349－2354.
14) **UK Prospective Diabetes Study (UKPDS) Group**：Lancet 1998；**352**：837－853.
15) **Howard AA, et al.**：Ann Intern Med 2004；**140**：211－219.
16) **Sone H, et al.**：Ann Intern Med 2004；**141**：408－409.
17) **Sone H, et al.**：Diabetes Care 2005；**28**：1463－1471.
18) **Sone H, et al.**：Diabetes Care 2006；**29**：145－147.

CLINICAL TRIALS ARTICLE Clinical Trials 2007; 4: 318–328

Estimation of treatment effect adjusting for dependent censoring using the IPCW method: an application to a large primary prevention study for coronary events (MEGA study)

Mizuki Yoshida[a], Yutaka Matsuyama[b] and Yasuo Ohashi[b], for the MEGA Study Group

Background The MEGA study is a randomized controlled trial conducted in Japan to evaluate the primary preventive effect of pravastatin against coronary heart disease (CHD), in which 8214 subjects are randomized to diet or diet plus pravastatin. Pravastatin reduces the incidence of CHD (hazard ratio = 0.67; 95%CI: 0.49–0.91). In the MEGA study, in addition to the usual loss to follow-up cases, there is another problem of drop-outs due to the refusal of further follow-up at 5 years.

Purpose To estimate the treatment effect adjusting for some types of dependent censorings observed in the MEGA study and to assess the sensitivity of standard analysis results for these censoring cases.

Methods The proposed method is a straightforward extension of the inverse probability of censoring weighted (IPCW) method for settings with more than one reason for censoring, where the propensities for drop-outs are modeled separately for each reason. Simulation studies are also conducted to compare the properties of the IPCW estimate with the standard analysis assuming independent censorings.

Results Simulation studies show that the IPCW estimate can correct for selection bias due to dependent censoring that can be explained by measured factors, while the standard analysis is biased. Applying the proposed method to the MEGA study data, several prognostic factors are associated with the censoring processes, and after adjusting for these dependent censorings, slightly larger treatment effects for pravastatin are observed for both CHD (primary endpoint) and stroke (secondary endpoint) events.

Limitations The method developed is based on the fundamental assumption of sequentially ignorable censoring.

Conclusions Our proposed method provides a valuable approach for estimating treatment effect adjusting for several types of dependent censorings. Dependent censorings observed in the MEGA study did not cause a severe selection bias attributable to the covariates and the results from the standard analysis were robust in relation to the censorings. *Clinical Trials* 2007; 4: 318–328. http://ctj.sagepub.com

[a]Biometrics Division, The Clinical Service Provider, EPS Co., Ltd., 2-3-19 Koraku, Bunkyo-ku, Tokyo 112-0004, Japan
[b]Department of Biostatistics/Epidemiology and Preventive Health Sciences, School of Health Sciences and Nursing, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
Author for correspondence: Yutaka Matsuyama, Department of Biostatistics, School of Health Sciences and Nursing, University of Tokyo, 7-3-1 Hongo Bunkyo-ku, Tokyo 113-0033, Japan, Tel: +81-3-5841-3519, Fax: +81-3-5841-3527. E-mail: matuyama@epistat.m.u-tokyo.ac.jp
Grant Support: This research was supported in part by Grant-in-Aid for Scientific Research (A) No. 16200022. Research funds for MEGA Study were provided by the Japanese Ministry of Health, Labor and Welfare for the first 2 years of the study, and thereafter the study was funded by Sankyo Co Ltd, Tokyo.

10.1177/1740774507081224

Introduction

The management of elevated cholesterol in the primary prevention group of adult Japanese (MEGA) study is a randomized controlled trial conducted in Japan to evaluate the primary preventive effect of a statin against coronary heart disease (CHD) in daily clinical practice [1]. In this prospective, randomized, open-labeled, blinded-endpoints (PROBE) design study, men and postmenopausal women aged 40–70 years with hypercholesterolemia (total cholesterol (TC) level: 220–270 (mg/dL)) and no history of CHD or stroke were randomized to diet (diet group) or diet plus pravastatin 10–20 mg daily (pravastatin group). The primary endpoint was the first occurrence of CHD, comprising fatal and nonfatal myocardial infarction, angina, cardiac and sudden death, and a coronary revascularization procedure. One of the secondary endpoints was the first occurrence of stroke.

Between February 1994 and March 1999, a total of 15 210 persons visiting an outpatient clinic were registered throughout Japan. Of the 15 210 subjects who met the inclusion criteria regardless of their TC level and who provided signed informed consent, 8214 who met the TC criterion were randomized to either diet or diet plus pravastatin treatment using the permuted block method with stratification according to gender, age, and medical institution. Of the randomized subjects, 382 were excluded; 94 withdrew their consent, 224 had exclusion criteria violation after randomization, and 64 had no recorded data after randomization. The remaining 7832 patients were analyzed (3966 diet group; 3866 pravastatin group). Follow-up was continued until March 2004. The incidence of CHD was significantly lower by 33% in the pravastatin group than in the diet group (hazard ratio = 0.67; 95%CI: 0.49–0.91; $p = 0.01$) [2].

A common problem encountered in any survival analysis is censoring data with a possible non-ignorable response mechanism. The response mechanism, which is the reason whether or not a response is obtained, is said to be non-ignorable if it depends on a subject's unobserved response [3]. If the censoring times are stochastically independent of survival time, the censoring is ignorable and standard survival analysis methods assuming independent censoring is valid. For example, an end-of-study censoring is completely determined by the enrollment time. If the survival does not change over time, such censoring time is independent of survival time. The assumption of independence, however, can never be verified from observed data and often may not be justified in practical settings. For example, one would suspect that drop-out subjects are different from the other subjects with respect to many background characteristics including the histories of prognostic factors such as lipid values. This type of drop-out may be dependent on the event of interest. The Kaplan–Meier estimator or the log-rank test under the assumption of independence will be inconsistent in the presence of dependent censorings [4].

In the MEGA study, although the follow-up period was initially scheduled for 5 years, based on the recommendation of the Data and Safety Monitoring Committee, the study was continued an additional 5 years to increase the number of events, and thus, patients who provided written consent at 5 years to continue the study were followed until the end of March 2004 [1,2]. Therefore, in addition to the usual loss to follow-up cases, there is another problem of drop-outs due to the refusal of further follow-up at 5 years. To ensure that the results in MEGA study are robust in relation to its censorings, it is important to assess the sensitivity of standard analysis results [2] to these possibly dependent drop-out cases.

Recently, Robins and colleagues proposed the inverse probability of censoring weighted (IPCW) method for the analysis of data with informative censoring [5–9]. The underlying idea of the IPCW method is to base estimation on the observed responses but weight them to account for the probability of remaining in the study. The propensities for drop-outs can be estimated as a function of the observed responses prior to drop-outs, and also as a function of the covariates and any additional variables or subject characteristics that are thought likely to predict drop-outs. The IPCW method can be used to correct for bias due to dependent censoring when the dependent censoring can be explained by measured prognostic factors.

In this article, we extend the IPCW approaches for time-to-event data [7] to settings with more than one reason for censoring. To obtain the probability of remaining in the study, we use separate models for each drop-out process, because, in the MEGA study, one type of drop-out dominates at 5 years, and the other type dominates otherwise. This modeling strategy is important, because there is a possibility of important differences in the effect of various predictors on each separate type of drop-out, and causal interpretation of the IPCW estimates depends on the correct specification of the model for drop-out [7,9]. For the usual drop-out cases such as loss to follow-up, the cause-specific hazards of censoring are modeled by the time-dependent Cox proportional hazards model, in which the treatment group-specific baseline hazard and parameters are separately assumed in the two treatment groups. For the

drop-outs due to the refusal of follow-up, the probability of drop-outs at 5 years is modeled by the logistic regression model. These two estimated weights are combined in order to construct the IPCW Kaplan–Meier estimator and the IPCW log-rank statistic. The remainder of this article is divided into five sections: presenting the MEGA study data; describing the proposed IPCW methodology; presenting the simulation studies to evaluate the performance of the IPCW estimation method; presenting the analysis results of the MEGA study data; and finally, conclusion with some discussion.

MEGA study

We will briefly describe the MEGA study data. Full details on the design, conduct, and main clinical results have been reported [1,2]. Table 1 shows the baseline characteristics of the analyzed patients. There was no clinical difference between the two groups in baseline characteristics. Women accounted for 68.4% (5356 patients) of the study population. Mean body mass index (BMI) was 23.8 (kg/m^2). Mean TC, low-density lipoprotein cholesterol (LDL-C), and high-density lipoprotein cholesterol (HDL-C) levels were 242.6, 156.6, and 57.5 (mg/dL), respectively. Median triglyceride (TG) level was 127.5 (mg/dL). Of the study patients, 41.8 and 20.8% had hypertension and diabetes mellitus based on physician diagnosis, respectively.

After randomization, patients were followed at months 1, 3, and 6 and thereafter every 6 months. At each visit, data on treatment compliance, use of concomitant drugs, onset of events, occurrence of adverse events, and laboratory tests including serum lipids were collected by the investigators. Additionally, an ECG was obtained and evaluated annually. All endpoints were reviewed strictly by the blinded Endpoint Committee and additional information obtained from the physician as needed [1]. A total of 7832 patients were followed by 2658 physicians in 1320 hospitals. The follow-up period was 41195 person-years (mean follow-up period 5.3 years). Table 2 shows the types and numbers

Table 1 Baseline characteristics of analyzed 7832 patients

Characteristics	Diet group $N = 3966$		Diet + pravastatin group $N = 3866$	
Age (years), mean (SD)	58.4	(7.2)	58.2	(7.3)
Women, No. (%)	2718	(68.5)	2638	(68.2)
BMI (kg/m^2), mean (SD)	23.8	(3.0)	23.8	(3.1)
Current smoker, No. (%)	572	(14.4)	612	(15.8)
Current drinking, No. (%)	1183	(29.8)	1180	(30.5)
Hypercholesterolemia medication history, No. (%)	621	(15.7)	586	(15.2)
Hypertension, No. (%)	1664	(42.0)	1613	(41.7)
Diabetes, No. (%)	828	(20.9)	804	(20.8)
TC (mg/dL), mean (SD)	242.6	(12.1)	242.6	(12.0)
G (mg/dL), median (inter-quartile range)	127.5	(95.0–179.0)	127.4	(95.7–176.5)
HDL-C (mg/dL), mean (SD)	57.5	(15.1)	57.5	(14.8)
LDL-C (mg/dL), mean (SD)	156.5	(17.3)	156.7	(17.6)

SD: standard deviation; BMI: body mass index; TC: total cholesterol; TG: triglyceride; HDL-C: high-density lipoprotein cholesterol; LDL-C: low-density lipoprotein cholesterol

Table 2 Types and numbers of events

Types of events	Diet group		Diet + pravastatin group	
	Number	(%)	Number	(%)
CHD	101	(2.5)	66	(1.7)
Loss to follow-up	546	(13.8)	594	(15.4)
Refusal of follow-up by patients	278	(7.0)	270	(7.0)
Refusal of follow-up by institutions	165	(4.2)	162	(4.2)
End-of-study censoring	2876	(72.5)	2774	(71.8)
Total	3966	(100)	3866	(100)

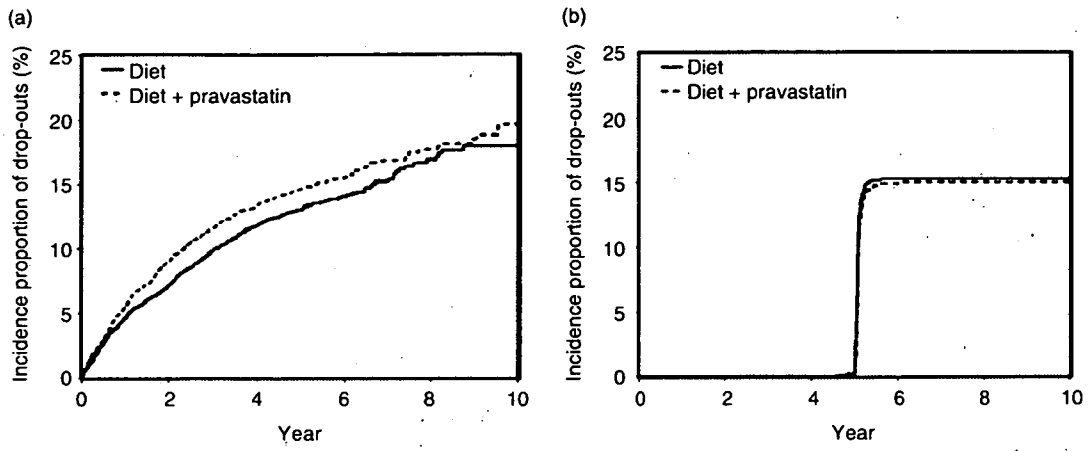


Figure 1 Incidence proportion for drop-outs (a) loss to follow-up; (b) refusal of follow-up by patients

of events in each treatment group. The events were divided into five categories: 1. CHD events; 2. loss to follow-up; 3. refusal of follow-up at 5 years by patients; 4. refusal of follow-up at 5 years by institutions; and 5. no events at the end of study. The withdrawal of informed consent occurring except at 5 years was included in the category of loss to follow-up (283 patients in diet group, 382 patients in pravastatin group). The refusal of follow-up at 5 years was divided into two categories, because, when obtaining the consent to continue the study, each Institutional Review Board (IRB) firstly made the decision regardless of the patient's intention. This institution-specific drop-out at 5 years was not related to the patient's medical histories and was thought to be an end-of-study censoring. These censorings at the end of study (refusal of follow-up by institutions and no events at the end of study) were not considered dependent censoring, because there was a fixed known calendar date at which the follow-up ended. Therefore, the second (Reason 1: loss to follow-up) and the third (Reason 2: refusal of follow-up by patients) categories were regarded as dependent censorings in this study.

Figure 1 shows the Kaplan–Meier curves for the event of drop-outs, Reasons 1 and 2, respectively. For Reason 1, more drop-outs were observed in the pravastatin group. For Reason 2, the times of drop-outs were distributed around 5 years after the start of follow-up and there was no distributional difference between the two treatment groups. Figure 2 shows the Kaplan–Meier curves for the CHD events that censored all drop-out cases at their event times. The incidence of CHD was significantly lower by 33% in the pravastatin group than in the diet group (hazard ratio = 0.67; 95%CI: 0.49–0.91; $p = 0.01$) [2].

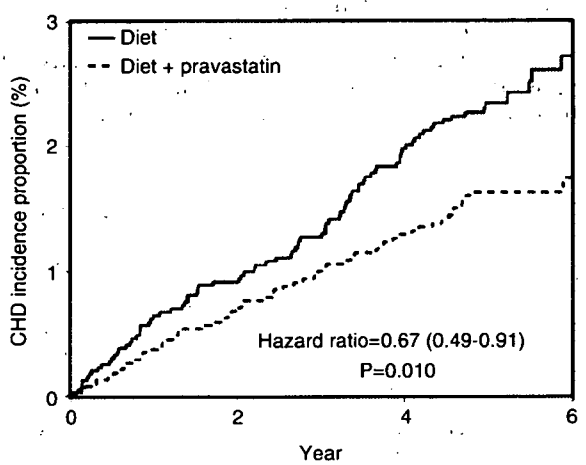


Figure 2 Incidence proportion for CHD events under the assumption of independent censoring

IPCW methods

Notation and assumption of no unmeasured confounders for censoring

Let T_i and C_i be the potential failure (occurrence of CHD events) time and the potential censoring time for subject i $(i = 1, \ldots, n)$, respectively. C_i is the minimum of C_{ij} $(j = 1, 2, 3)$, where C_{i1} denotes a censoring time of loss to follow-up (Reason 1), C_{i2} denotes a censoring time due to refusal of follow-up by patients (Reason 2), and C_{i3} denotes a censoring time at the end of study. The observable data are n i.i.d. copies of $X = \min(T, C_1, C_2, C_3)$, type of event J ($j = 0$, if CHD events are observed), treatment group indicator variable R ($R = 1$ if diet plus pravastatin group, and $R = 0$ if diet group), and

covariate history $\overline{V}_X$, where $\overline{V}_t = \{V_s: 0 \le s \le t\}$, and V_s is a vector of possibly time-dependent prognostic factors for T recorded at time s.

In order to identify the survival time in the presence of dependent censoring, we assume the following relation in the censoring process,

$$\lambda_{C_j}(t|R, \overline{V}_t, T, T > t) = \lambda_{C_j}(t|R, \overline{V}_t, T > t), \quad (1)$$

where $j = 1, 2$ and $\lambda_{C_j}(t|(\cdot), T > t)$ is the cause-specific hazard of censoring at time t given both $X = \min(T, C_1, C_2, C_3)$ exceeds at t and the information in $(\cdot)$. This assumption means that, conditional on the treatment group and on the recorded history until time t, the cause-specific hazard of censoring C_j $(j = 1, 2)$ at time t does not further depend on the possibly unobserved CHD event time T. This fundamental assumption is called 'no unmeasured confounders for censoring' [10] and is equivalent to a sequential version of Rosenbaum and Rubin's strong ignorability assumption [11]. The assumption specifies that, among subjects with the same recorded past, the population of subjects censored due to each specific cause at time t has the same distribution of the outcome of interest as that of the population of uncensored subjects at time t. The assumption will be satisfied, in particular, when the censoring process is ignorable or missing at random (MAR) in the terminology of missing data analysis [3]. In practice, we would not expect this assumption to be precisely true, but given a rich collection of prognostic factors recorded in $\overline{V}_t$, it may well be approximately true.

Estimation of the probability of remaining in the study (Reason 1)

The IPCW approach is to artificially regard subjects as dependently censored at the first time a subject was censored by either loss to follow-up or refusal of follow-up. To correct for dependent censoring, we need to estimate the treatment group-specific hazards of censoring conditional on time-dependent prognostic factors for CHD [7]. For the drop-outs due to loss to follow-up (Reason 1), the time-dependent Cox proportional hazards model for censoring is used for the right-hand side of Equation (1),

$$\lambda_{C_1}(t|R, \overline{V}_t, T > t) = \lambda_{0R}(t) \exp(\alpha_R \overline{V}_t), \quad (2)$$

where the treatment group-specific baseline hazard $\lambda_{0R}(t)$ and the treatment group-specific regression parameters α_R are assumed, because both the baseline hazard and covariate effects may depend on treatment group. For estimating the hazard of censoring (2) conditional on covariates, CHD events and other censoring types are censored at their event times.

Under the assumption of no unmeasured confounders for censoring (1) and the proportional hazards model for cause-specific hazards of censoring (2), the conditional probability of being uncensored due to the Reason 1 for subject i is provided by the following time-dependent extension of the Kaplan–Meier estimator,

$$\hat{K}_{i1}(t) = \prod_{u: X_u < t, \sigma_{u1} = 1, R_u = R_i} \exp\left[-\hat{\lambda}_{0R}(X_u) \exp(\hat{\alpha}_R \overline{V}_{iX_u})\right], \quad (3)$$

where $\hat{\lambda}_{0R}(X_u) = \sigma_{u1} / \sum_{i=1}^{n} \exp(\hat{\alpha}_R \overline{V}_{iX_u}) Y_i(X_u) I \times (R_i = R)$ is the Breslow estimator of the baseline hazard function for censoring $j = 1$ in treatment group R, and $Y_i(t)$ takes the value of one if subject i is at risk at time t, and zero otherwise. σ_{u1} takes the value of one if the subject is censored for Reason 1, and zero otherwise. For any proposition A, $I(A)$ equals one if A is true and zero otherwise.

Estimation of the probability of remaining in the study (Reason 2)

For the drop-outs due to refusal of follow-up by patients (Reason 2), because the drop-out times are fixed at 5 years, the probability of drop-outs is modeled by the following logistic regression model,

$$\text{logit } \Pr(D_i = 1|R, \overline{V}_5, Z_i = 1) = \gamma_{0R} + \gamma_R \overline{V}_5, \quad (4)$$

where D_i takes the value of one if subject i refuses further follow-ups at 5 years, and zero otherwise, Z_i takes the value of one if subject i experiences the re-informed consent at 5 years, zero otherwise, and γ_{0R} and γ_R are the treatment group-specific regression parameters. Under the assumption of no unmeasured confounders for censoring (1) and the model (4), the conditional probability of being uncensored due to Reason 2 for subject i is estimated by

$$\hat{K}_{i2}(5) = 1 - \hat{Pr}(D_i = 1|R, \overline{V}_5, Z_i = 1). \quad (5)$$

Estimation of the IPCW survival function

The IPCW estimator is different from the ordinary estimator by weighting the contribution of a subject at risk by the inverse of the conditional probability of having remained uncensored. Using the above estimators of uncensored probability, $\hat{K}_{i1}(t)$ and $\hat{K}_{i2}(5)$, the contribution of a subject at risk at time t is weighted by the inverse of an estimate of

the conditional probability of having remained uncensored for both reasons until time t,

$$\hat{W}_i(t) = \begin{cases} \dfrac{1}{\hat{K}_{i1}(t)} & \text{for } t < 5 \\ \left(\dfrac{1}{\hat{K}_{i1}(t)}\right) \times \left(\dfrac{1}{\hat{K}_{i2}(5)}\right) & \text{for } t \geq 5 \text{ and } Z_i = 1. \\ \dfrac{1}{\hat{K}_{i1}(t)} & \text{for } t \geq 5 \text{ and } Z_i = 0 \end{cases}$$

Here, we assume that the conditional probabilities are bounded away from zero with probability 1 for each subject i, that is, $\hat{K}_{ij}(t) > 0$. This assumption will be satisfied unless their conditional probabilities are structural zero, that is, $\hat{K}_{ij}(t) = 0$ for some values of $\bar{V}_t$. Under this assumption, the IPCW Kaplan–Meier estimator of the treatment group-specific survival of not having CHD events through time t is

$$\hat{S}_T(t|R) = \prod_{\{i:X_i<t\}} \left\{ 1 - \frac{\delta_i \hat{W}_i(X_i) I(R_i = R)}{\sum_{u=1}^{n} Y_u(X_i) \hat{W}_u(X_i) I(R_u = R)} \right\} \tag{6}$$

where δ_i is the failure time indicator that takes the value of one if the subject failed and zero if the subject is censored. This IPCW Kaplan–Meier estimator for CHD events in treatment group R differs from the ordinary Kaplan–Meier estimator in that the contribution of a subject at any time X_i is weighted by the subject-specific weight $\hat{W}_i(X_i)$. In the IPCW estimator (6), the quantity, $\delta_i \hat{W}_i(X_i) I(R_i = R)$, estimates the number of subjects in treatment group R who would have been observed to fail at time X_i in the absence of drop-outs, while the quantity, $\sum_{u=1}^{n} Y_u(X_i) \hat{W}_u(X_i) I \times (R_u = R)$, estimates the number of subjects who would have been alive and at risk at time X_i in the absence of drop-outs. Thus, the ratio estimates the hazard of CHD event at X_i in the absence of drop-outs; it follows that (6) estimates the probability $S_T(t|R)$ of surviving without failure (i.e., of remaining CHD-free) until time t in the absence of drop-outs. Under assumption (1) and the correct specification of weights, Robins [5] proves that under mild regularity conditions, the IPCW estimator (6) gives a consistent estimator of our target causal estimand $S_T(t|R)$. Inverse probability weighted estimators have been previously considered by Horvitz and Thompson [12] in the sample survey literature. Satten and Datta [13] give an elementary discussion of the IPCW estimators.

Comparison of the IPCW survival function

We used the Cox proportional hazards model to compare the IPCW survival distribution between the two treatment groups. The model is

$$\lambda_T(t|R) = \lambda_0(t) \exp(\beta R)$$

where $\lambda_T(t|R)$ is the potential hazard of CHD events at time t in the treatment group R. The IPCW Cox partial likelihood score $U(\beta)$ for β differs from the ordinary Cox partial likelihood score in that the contribution of the subject u at risk at time X_i is weighted by $\hat{W}_u(X_i)$, that is,

$$U(\beta) = \sum_{i=1}^{n} \delta_i \hat{W}_i(X_i) \times \left\{ R_i - \frac{\sum_{u=1}^{n} Y_u(X_i) \hat{W}_u(X_i) R_u \exp(\beta R_u)}{\sum_{u=1}^{n} Y_u(X_i) \hat{W}_u(X_i) \exp(\beta R_u)} \right\} \tag{8}$$

Under the assumption (1) and the correct specification of weights, Robins [5] proves that under mild regularity conditions, the weighted estimating equations $U(\beta) = 0$ gives a consistent and asymptotically normal estimator of the parameter β, which can be interpreted as the treatment effect in the absence of drop-outs.

The use of individual weights induces within-subject correlation and we must take this correlation into consideration in the calculation of variance. In the calculation of a confidence interval, we used the robust variance estimate [14]. It provides a conservative confidence interval for the parameter of interest, that is, the 95% Wald confidence interval calculated as $\beta \pm 1.96 \times$ (robust standard error), which is guaranteed to cover the true value of β at least 95% of the time in large samples [14,15].

Simulation studies

Settings of simulations

To evaluate the performance of the IPCW estimation method, we carried out simulation studies under three conditions: missing completely at random (MCAR), missing at random (MAR), and missing not at random (MNAR). We simulated data from two treatment groups, coded as $R = 0$ (control treatment) or $R = 1$ (test treatment). About equal sample size of 500 for each group was randomly generated (total sample size was 1000). The simulations were based on 1000 replications, so that the estimated coverage probability of a true

95% confidence interval would have a simulation accuracy of ≈1.35%.

For each subject i ($i=1,\ldots,1000$), a time-dependent covariate L_{it} ($t=0,\ldots,4$) was generated via the following mixed effect model,

$$L_{it} = 10 - 0.10 \times I(R_i = 0) \times t - 1.70 \times I(R_i = 1) \times t + b_{0i} + b_{1i} \times t + \varepsilon_{it}.$$

The random effects b_{0i} and b_{1i} were generated from a bivariate normal distribution with means of zero and their variance of 3.0 and 2.5, respectively, with the correlation coefficient of 0.8. The random error ε_{it} was generated from a normal distribution with a mean of zero and a variance of 0.8. For each subject, L_{it} was supposed to be observed until just before the observed failure time.

The potential failure time T_i was generated from the following exponential distribution with hazard λ,

$$S(t) = \exp(-\lambda t),$$

where $\lambda = \exp(\alpha_0 + \alpha_1 R_i + \alpha_2 L_{i0})$, $(\alpha_1, \alpha_2) = (-0.5, 0.3)$, and $\alpha_0 = -6.0$ (MCAR and MAR cases), $\alpha_0 = -5.7$ (MNAR case). The drop-outs were assumed to occur at four time points ($t=1,\ldots,4$) and censoring at the end of follow-up ($t=5$) was considered independent censoring. For simplicity, only one type of drop-out was considered, where the drop-out indicator variable D_{it} ($D_{it}=1$ if drop-outs, $D_{it}=0$ if otherwise) was generated from the following conditional model,

$$\text{logit } \Pr(D_{it} = 1 | D_{i(t-1)} = 0, L_{i(t-1)}, t, T_i > t) = \beta_0 + \beta_1 t + \beta_2 L_{i(t-1)} + \beta_3 T_i, \quad (9)$$

where $t=1,\ldots,4$, $\beta_2=\beta_3=0$ corresponds to MCAR case (β_0 and β_1 were set to be -2.0 and -0.4, respectively), $\beta_2 \neq 0$, $\beta_3=0$ corresponds to MAR case (β_0, β_1, and β_2 were set to be -6.0, -0.4, and 0.4, respectively), and $\beta_3 \neq 0$ corresponds to MNAR case (β_0, β_1, β_2, and β_3 were set to be -3.5, -0.4, 0.4, and -0.3, respectively). In the above settings, although the potential failure time T_i is assumed to be directly dependent on group and baseline covariate L_{i0}, because of the high correlation between L_{i0} and L_{it}, the larger the values of L_{it}, the T_i is shorter and the more drop-out cases will be observed. The percentages of event, drop-outs, and censoring at the end of follow-up are roughly 10, 20, and 70% ($R=1$), 20, 20, and 60% ($R=0$), respectively. The observed failure time X_i was set to $X_i=t$ for drop-out cases at time t, $X_i=T_i$ for event cases whose potential failure time is not exceeding 5, and $X_i=5$ for censoring cases at the end of follow-up whose potential failure time is >5.

In each repetition of simulations, the proportional hazards model including a group variable R as a covariate was fitted to the observed failure time X_i and the estimate of the log(hazard ratio), $\hat{\theta}_s$ ($s=1,\ldots,1000$), was calculated. The following three proportional hazards models were fitted. The first one was the standard analysis ignoring the time-dependent covariate L_{it}, where all drop-out cases were assumed to be censored at their drop-out times (assumption of independent censoring). The second one was the adjusted analysis including the time-dependent covariate L_{it} as covariates under the assumption of independent censoring. The third one was the proposed IPCW analysis, where the weights were estimated by fitting the model (9) with $\beta_3=0$ to the observed data.

The result from the analysis for data that had been observed in the absence of drop-outs was regarded as a true value of the log(hazard ratio) in each repetition. The observed failure time that had been observed in the absence of drop-outs A_i was defined to be $A_i=T_i$ for event cases whose potential failure time is not exceeding 5, and $A_i=5$ for the censoring cases at the end of follow-up whose potential failure time is >5.

Results of simulations

Simulations were evaluated in terms of the percent relative bias, mean squared error (MSE), and coverage probability of nominal 95% large sample confidence intervals for the estimate of the log(hazard ratio) for group effect. The percent relative bias was computed as $(1/1000)\sum(\hat{\theta}_s - \bar{\theta})/\bar{\theta} \times 100$, where $\hat{\theta}_s$ is the estimate of $\bar{\theta}$ (average of true values for the log(hazard ratio)) from the sth simulated replication.

Table 3 shows the results. Under the MCAR setting, both estimates from the standard and the IPCW analysis were nearly unbiased and their coverage probabilities were close to the nominal level of 95%, while the adjusted estimate was largely biased with anticonservative coverage probability. Under the MAR setting, as expected based on the theory, the selection bias due to the time-dependent covariate L_{it} was adjusted by the IPCW analysis, while the estimate from the standard analysis underestimated the treatment effect because the subjects with larger values of L_{it} and shorter potential failure time tended to drop out. MSE from the IPCW analysis was slightly larger than that of the standard one. Under the MNAR setting, the IPCW analysis could not adjust the selection bias due to the violation of the assumption of no unmeasured confounders for censoring,

Table 3 The results of simulations for the estimate of treatment effect

	Estimation method	True value	Relative bias (%)	MSE	95% coverage
MCAR	Standard	−0.485	0.60	0.023	95.0
	Adjusted		−44.23	0.072	71.0
	IPCW		0.62	0.023	95.1
MAR	Standard	−0.483	−3.20	0.027	94.2
	Adjusted		−43.43	0.072	73.2
	IPCW		−0.57	0.032	94.9
MNAR	Standard	−0.479	−3.31	0.025	94.6
	Adjusted		−29.57	0.045	84.2
	IPCW		−2.46	0.026	95.3

MSE: mean squared error
Standard analysis is the proportional hazards model including only a group variable as a covariate under the assumption of independent censoring. Adjusted analysis is the same model except for including the time-dependent covariates.

although the bias was slightly smaller than that of the standard one.

Analysis of MEGA study data

Factors affecting each drop-out

To construct the IPCW estimators, it is necessary to estimate the subject-specific weight $\hat{W}_i(X_i)$ conditional on time-dependent prognostic factors for failure. We have to choose variables for modeling the censoring process so as to make assumption (1) plausible. As causal interpretation of estimates depends on the correctness of (1), making the censoring process ignorable is more important than fitting a parsimonious model. However, because of the large number of potential prognostic factors included in $\overline{V}_t$, it may be useful to reduce these to a relevant subset. In general, for a time-dependent prognostic factor to cause selection bias or confounding, it must be a prognostic factor for both failure and censoring.

To estimate the subject-specific weight $\hat{W}_i(X_i)$, we used five time-dependent factors as well as twelve baseline factors shown in Table 1. Among baseline factors, missing data were observed in the values of BMI (0.24%), current smoking (0.18%), and drinking (0.17%). The missing values of BMI were imputed by the mean value of 23.8 (kg/m^2). The later two factors were imputed by zero (no smoking and no drinking, respectively). Five time-dependent factors were four lipids (TC, TG, HDL-C, and LDL-C) and treatment actually received. For the missing data of lipid values (21.5%), the regression imputations were separately conducted, where 12 baseline factors, allocation group, and the last observed lipid value were included as covariates in each prediction model. For the missing data of treatment actually received (10.5%), the last observation carried forward method was used to impute the missing values.

To estimate the treatment group-specific hazards of censoring due to Reason 1 and the probability of drop-outs at 5 years due to Reason 2, five combinations of covariates $\overline{V}_t$ were used in both Models (2) and (4), accounting for the multicollinearity of covariates. First one included all five time-dependent factors and twelve baseline factors as covariates (Model 1). Second one excluded all TC values and time-dependent treatment group from Model 1 (Model 2). Third one excluded all baseline lipid values from Model 2 (Model 3). Fourth one excluded all TC values, time-dependent TG, HDL-C, and LDL-C from Model 1 (Model 4). Last one included only significant variables in the Model 2 (Model 5). The values of TG and HDL-C were included into the above five models by taking their logarithm. For the five time-dependent factors, the most recent recorded values were included as covariates in the prediction model.

Table 4 shows the effect estimates of each factor associated with two types of censorings. The results from Model 3 are presented, because the results from the other models were similar to those shown in Table 4, except previous non-use of pravastatin also predicted drop-outs from the study. For Reason 1, patients who have hypertension, diabetes mellitus, or hypercholesterolemia medication history tended to remain in the study. In the pravastatin group, the higher the values of TG or LDL-C during the study period, the more drop-out cases that were observed. For Reason 2, patients with hypertension or lower values of TG or HDL-C during the study period were likely to consent to the further follow-up at 5 years.

Table 4 Factors affecting each drop-out (results from Model 3)

	Loss to follow-up (Reason 1)				Refusal of follow-up by patients (Reason 2)			
	Diet		Diet + pravastatin		Diet		Diet + pravastatin	
Factors	HR	95% CI	HR	95% CI	OR	95% CI	OR	95% CI
Baseline								
Age (years)	1.01	0.99, 1.02	1.00	0.99, 1.02	0.99	0.97, 1.01	1.00	0.98, 1.02
Women	1.08	0.85, 1.37	1.00	0.80, 1.27	0.80	0.55, 1.18	1.28	0.86, 1.89
BMI (kg/m^2)	1.01	0.98, 1.04	0.98	0.95, 1.01	1.10	0.97, 1.06	0.98	0.94, 1.03
Current smoker	1.13	0.88, 1.44	1.21	0.94, 1.54	1.14	0.75, 1.72	1.23	0.81, 1.85
Current drinking	1.14	0.91, 1.45	1.03	0.83, 1.28	0.81	0.55, 1.17	1.17	0.80, 1.70
Medication history	0.84	0.66, 1.08	0.63	0.48, 0.81	0.87	0.58, 1.29	0.70	0.45, 1.08
Hypertension	0.82	0.68, 0.98	0.91	0.77, 1.07	0.79	0.60, 1.05	0.76	0.57, 1.01
Diabetes	1.02	0.82, 1.25	0.72	0.58, 0.90	0.83	0.60, 1.18	1.01	0.73, 1.42
Time-dependent								
TG (mg/dL)	1.11	0.92, 1.35	1.30	1.08, 1.57	1.56	1.08, 2.25	1.72	1.20, 2.48
HDL-C (mg/dL)	0.82	0.54, 1.26	1.11	0.74, 1.67	3.14	1.54, 6.41	1.76	0.84, 3.69
LDL-C (mg/dL)	1.00	0.99, 1.01	1.01	1.01, 1.01	1.00	1.00, 1.01	1.00	1.00, 1.01

HR: hazard ratio; OR; odds ratio; CI: confidence interval; medication history: hypercholesterolemia medication history

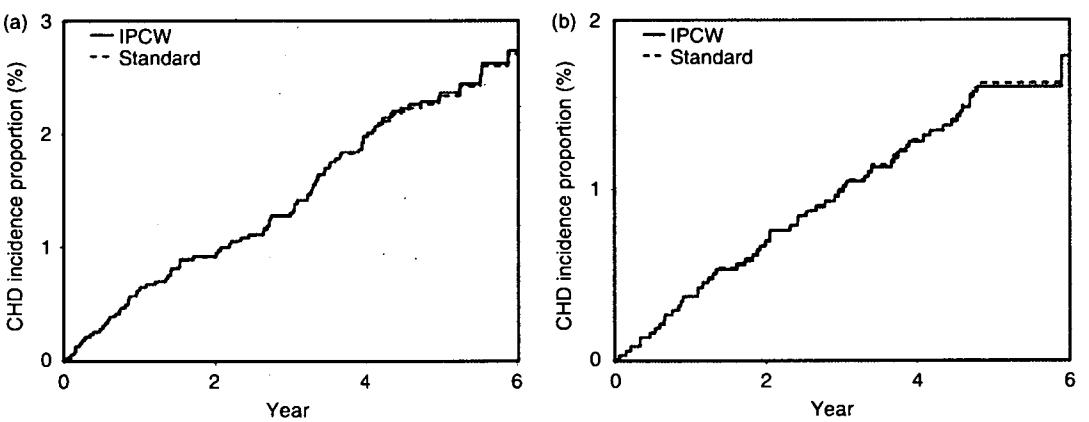


Figure 3 IPCW incidence proportion for CHD events in each treatment group. In each group, the solid line is the IPCW estimate and the dashed one is the standard estimate shown in Figure 2: (a) Diet group; (b) Diet + pravastatin group

Estimation of treatment effect adjusting for dependent censoring

Figure 3 shows the IPCW Kaplan–Meier curves for the CHD events in each treatment group. In each group, the solid line is the IPCW estimate whose weights were calculated from Model 1, and the dashed one is the standard estimate shown in Figure 2. In both treatment groups, the adjusted curves were almost the same as those obtained by assuming all drop-out cases as independent censoring. Table 5 shows the estimates of treatment effect under several models. Hazard ratios for stroke event, which was one of the secondary endpoints in the MEGA study, were also presented. Analysis models for stroke were the same as those for CHD events, and similar results for factors associated with censorings were observed (not shown) as shown in Table 4. For both CHD and stroke events, a slightly larger treatment effect was observed by the IPCW analysis. The IPCW estimates did not change under different models for the estimation of weights.

Discussion

In this article, we developed a method for the estimation of treatment effect adjusting for

Table 5 Estimates of treatment effect for CHD and stroke

		CHD			Stroke		
Method		HR	95% CI	*p*-value	HR	95% CI	*p*-value
Standard		0.67	0.49, 0.91	0.010	0.83	0.57, 1.21	0.33
IPCW	Model 1	0.65	0.48, 0.89	0.007	0.81	0.56, 1.18	0.27
	Model 2	0.66	0.48, 0.90	0.008	0.81	0.56, 1.18	0.28
	Model 3	0.66	0.49, 0.90	0.009	0.81	0.56, 1.17	0.26
	Model 4	0.66	0.49, 0.91	0.009	0.81	0.56, 1.18	0.27
	Model 5	0.66	0.48, 0.90	0.008	0.82	0.57, 1.19	0.29

HR: hazard ratio; CI: confidence interval; standard method is the analysis assuming all types of censorings as independent.

dependent censoring using the IPCW approach. This proposed method is a straightforward extension of Robins and Finkelstein [7] method for settings with two or more reasons for censoring. In real clinical trials, there are several different types of reasons for censoring and it is likely that each process has its own reason, that is, the covariates history through each censoring. Our proposed approach was a relatively easy method for accounting for the differences in the reasons for censoring by estimating the weights separately in the framework of the IPCW methodology.

It is important to note that our results are based on the fundamental assumption (1) that the cause-specific hazard of censoring can be totally explained by the treatment group and the recorded history of the covariates. This assumption is a non-identifiable assumption and is not testable from the observed data. There is a possibility of a residual effect due to unmeasured prognostic factors for censoring. However, in the MEGA study, many clinically important prognostic factors were measured and all of them were used as covariates to predict the probability of remaining in the study. In addition to the five prediction models shown in Table 5, the analyses based on other prediction models, in which time-dependent covariates were entered in different ways, such as the difference from the baseline or the absolute past two values, were conducted, and the IPCW estimates were shown to be insensitive to the selection of the prediction models conditional on the measured covariates. Therefore, a departure from the assumption (1) will be small in our IPCW estimates.

In this article, we regarded both the institution-specific drop-out at 5 years and the end-of-study censoring as independent censoring. However, there may be a possibility of correlation between such censorings and the prognosis [16]. We also conducted the IPCW analyses, where all types of censorings were considered as potentially dependent ones. In this analysis, censoring due to institutional refusal at 5 years was separately modeled by the logistic regression model such as (4), and the end-of-study censoring was modeled by specifying the cause specific hazard functions, where the time-dependent Cox proportional hazards models such as (2) were separately fitted for each cause of censoring. The IPCW hazard ratios from Model 3 in combinations of covariates $\overline{V}_t$ were 0.66 (95% CI: 0.48–0.90) for CHD events and 0.81 (95% CI: 0.56–1.18) for stroke event. Therefore, our informal assumption that the end-of study censoring including institutional refusal at 5 years was considered as independent censoring seemed to be reasonable.

In the analyses of the MEGA study, factors affecting drop-outs were different for each reason for censoring as well as for the treatment group. However, no history of medication for hypertension, diabetes mellitus, or hypercholesterolemia was related to both censorings, that is, patients with a relatively better medical condition tended to drop out. This suggested that troublesomeness or weak motivation for participating in the study might cause the drop-outs, because the MEGA study was a primary prevention study. Furthermore, the fact that patients with high HDL-C or not in the pravastatin group tended to drop-out may also explain the above possibility for the censoring process in the MEGA study. This censoring process was different from that observed in usual clinical trials where the occurrence of adverse events or deteriorating health condition are the primary reasons for study drop-out. On the other hand, for the time-dependent covariates, patients with higher values of TG tended to drop out, suggesting that non-compliance with the appropriate diet instructions or inadequate diet control during the study period were related to the censoring processes.

Compared with the standard analysis, a slightly larger treatment effect was observed by the IPCW analysis, but the difference was minimal.

In general, selection bias is a function of both the magnitude of censoring rate and how different the censored subjects are from uncensored ones in terms of prognosis. In the MEGA study, although the censoring proportions due to patient refusal of follow-up (Reason 2) were relatively small and there were no differences between treatment groups, those due to loss to follow-up (Reason 1) were relatively large and the differences were observed between treatment groups (Figure 1(a)). In the latter category of censoring, about half of the reasons for loss to follow-up were the withdrawal of informed consent (51.8% in diet group, 64.3% in pravastatin group). As shown in Table 4 and discussed previously, the censoring process observed in the MEGA study seemed to be unrelated to the occurrences of outcomes of interest. Therefore, the lack of effect of weighting in our data could be due to the fact that the weights are not highly related to the probabilities of disease outcome and thus would not have an appreciable effect in altering the point estimates. This result indicates that drop-outs observed in the MEGA study did not cause a severe selection bias attributable to the measured covariates and the standard analysis results [2] were robust for the drop-outs.

However, as shown in the results of the simulations, the estimate from the standard analysis is biased under the MAR setting where the drop-out process is dependent on the covariate histories. In many clinical trials, because we cannot safely say that the dependent censorings have not occurred, it is important to conduct the analysis accounting for the dependent censorings as well as the standard one and to compare their results. When their results differ remarkably, the reasons for drop-outs were examined in detail and the effects on the final conclusion in the clinical trial concerned should be discussed.

Acknowledgments

MEGA Study is supported by Sankyo Co. LTD. We also thank the referee, the associate editor, and the editor for their comments, which led to a much improved version of the article.

References

1. **Management of Elevated Cholesterol in the Primary Prevention Group of Adult Japanese (MEGA) Study Group.** Design and baseline characteristics of a study of primary prevention of coronary events with pravastatin among Japanese with mildly elevated cholesterol levels. *Circ J* 2004; **68**: 860–67.
2. **Nakamura H, Arakawa K, Itakura H *et al.*** Primary prevention of cardiovascular disease with pravastatin in Japan (MEGA Study): a prospective randomised controlled trial. *Lancet* 2006; **368**: 1155–63.
3. **Rubin DB.** Inference and missing data. *Biometrika* 1976; **63**: 581–92.
4. **Kalbfleish JD, Prentice RL.** *The Statistical Analysis of Failure Time Data* (2nd ed.) John Wiley, New York, 2002.
5. **Robins JM.** Information recovery and bias adjustment in proportional hazards regression analysis of randomized trials using surrogate makers. *American Statistical Association Proceedings of the Biopharmaceutical Section* 1993; 24–33.
6. **Robins JM, Rotnitzky A, Zhao LP.** Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. *J A Stat Assoc* 1995; **90**: 106–21.
7. **Robins JM, Finkelstein DH.** Correcting for noncompliance and dependent censoring in an AIDS clinical trial with inverse probability of censoring weighted (IPCW) log-rank tests. *Biometrics* 2000; **56**: 779–88.
8. **Scharfstein DO, Robins JM.** Estimation of the failure time distribution in the presence of informative censoring. *Biometrika* 2002; **89**: 617–34.
9. **Joffe MM, Hoover DR, Jacobson LP *et al.*** Estimating the effect of zidovudine on kaposi's sarcoma from observational data using a rank preserving structural failure-time model. *Stat Med* 1998; **17**: 1073–102.
10. **Robins JM.** Causal inference from complex longitudinal data. In Berkane M. Ed). *Latent variable modeling and application to causality* Springer-Verlag, New York, 69–117, 1997.
11. **Rosenbaum PR, Rubin DB.** The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983; 70: 41–55.
12. **Horvitz DG, Thompson DJ.** A generalization of sampling without replacement from a finite universe. *J A Stat Assoc* 1952; **47**: 663–85.
13. **Satten GA, Datta S.** The Kaplan-Meier estimator as an inverse-probability-of-censoring weighted average. *The American Statistician* 2001; **55**: 207–10.
14. **Lin DY, Wei LJ.** The robust inference for the Cox proportional hazards model. *J A Stat Assoc* 1989; **84**: 1074–78.
15. **Robins JM, Hernán MA, Brumback B.** Marginal structural models and causal inference in epidemiology. *Epidemiology* 2000; **11**: 550–60.
16. **Zhang J, Heitjan DF.** Nonignorable censoring in randomized clinical trials. *Clin Trials* 2005; **2**: 488–96.

臨床試験に対する生物統計学の社会的貢献：4半世紀の経験と今後の展望

Contribution of Biostatistics in Clinical Trials to Human Society: Experience of a Quarter of a Century and the Future Prospects

大橋靖雄
Yasuo Ohashi

東京大学医学系研究科 生物統計学
Department of Biostatistics, University of Tokyo
e-mail: ohashi@epistat.m.u-tokyo.ac.jp

1. はじめに

臨床医学の目的は，患者の疾患を正確に診断し適切に治療を行なうことにある．しかし，医療提供者側の知識・技術の不完全さ，患者・疾患の多様さに多く由来する治療効果の不確実さから，診断結果や治療結果には永遠に除去不可能な曖昧さが伴う．したがって，診断・治療によってもたらされるベネフィットと被るリスクは，ともに「可能性」として確率変数的性格を帯びる．これらに対する患者の重み付けも患者個々の価値判断を反映して異なるはずのものであり，診断・治療法の選択は，リスク・ベネフィット両者のバランスと資源の制約の中で，本来は充分な情報提示と理解，そして自発的意志を前提としたインフォームドコンセントによるべきである（とされる）．しかし，医療提供者と患者の有する情報の不均衡，医師・患者双方の意識の問題もあり，これまでの治療上の意志決定は，医師主導のパターナリズムの中で，曖昧な状況下で行われてきたといってよい．

近年の情報公開あるいは患者の権利主張の流れは，このような意志決定プロセスに大きな変革を与えようとしている．厚生労働省は，ここ数年，疾患毎に標準治療をまとめたガイドライン策定を各関連学会に依頼し，数多くの疾患ガイドラインやその案が発表されてきた．そして，このガイドライン策定過程で（医療関係者には周知であったが）改めて次の事態が浮き彫りになった．「わが国には客観的証拠 evidence が無い！」臨床医学系学会では，10 年ほど前に Evidence Based Medicine (EBM) という言葉が大流行したが，ここで evidence を提供するのが患者を対象とした臨床研究成果である．

一方，1993 年のソリブジン事件を重要な契機として，日本の臨床試験とくに治験の制度・それを取り巻く環境・そして内容は大きな変貌をとげた．制度の面では，現在の医薬品医療機器総合機構設立につながる審査体制の変革と 1998 年の新 Good Clinical Practice (GCP) 完全施行に代表される国際ハーモナイゼーション（以下略して ICH）の受け入れ，そして（まだ位置づけが十分には定まらないものの）医師主導治験制度の開始が象徴である．産業界は，これまでの官・産一体の護送船団方式から ICH-E5 に基づく海外データの受け入れ開始を経て，市場と開発の場を海外に求めるに至り，現象としては，国内治験の空洞化と治験グローバル化，そして相次ぐ企業合併