

連載

EBNのための統計の読み方⑮

生存時間解析法

一度起こると元に戻らない事象の解析

林 邦彦・高木廣文

生存時間解析法

一度起こると元に戻らない事象の解析

林 邦彦 HAYASHI Kunihiko 群馬大学医学部保健学科
 高木廣文 TAKAGI Hirofumi 新潟大学医学部保健学科

「きみの血液検査をしよう。そうすれば、答えが出る」「絶対確実な答え？」マックスが訊いた。ローガンは首を振った。……「この世で絶対確実なものはほんのわずかだよ、マックス」¹⁾

非可逆的事象とは

ショーン・コネリー扮する英国スパイ007が日本を舞台に活躍する映画のタイトルは、「007は二度死ぬ (You only live twice.)」だったが、通常、「人生は一度だけ (You only live once.)」だろう。死という事象は、007以外には皆に一度だけおとすれ、一度起きると元の状態には戻らない非可逆的事象と考えられる。このような事象で行う解析は、設定した時間的区間ごとにどのくらい起きるのかという時間経過での発生頻度や、いつ起きるのかという事象発生までの時間を検討する。その解析方法に、「生命表 (life table)」や「生存時間解析法 (survival analysis)」とよばれる手法がある。死という事象以外にも、患者の退院、免疫の獲得、最初の心筋梗塞発作、初恋、といったように一度だけしか起こらないと仮定してもよい事象には、これらの手法が

応用できる。

生命表

2000年の人口動態統計、国勢調査人口を用いて、わが国の男子の生命表を簡単に作成してみよう。①の左欄から、年齢 t 、その年齢の観察人口 S_t 、その年齢での1年間死亡数 d_t を表している。0歳児の1年あたりの死亡率 q_0 は、

観察開始時 年齢	1年間の観察 人口(人)*	死亡数 (人)	死亡率 (/年)	生存率 (/年)	累積生存率	累積死亡率
t	S_t	d_t	q_t (= d_t/S_t)	p_t (= $1-q_t$)	P_t (= $\prod p_t$)***	Q_t (= $1-P_t$)
0	612,148**	2107	0.00344	0.99656	0.99656	0.00344
1	598,688	330	0.00055	0.99945	0.99601	0.00399
2	603,618	200	0.00033	0.99967	0.99568	0.00432
3	609,138	160	0.00026	0.99974	0.99542	0.00458
4	607,410	136	0.00022	0.99978	0.99521	0.00479
5	611,277
.

2000年の人口動態統計、国勢調査人口のデータをもとに、男子について計算した。

* : 国勢調査での $t-1$ 歳人口と、 t 歳人口の平均値とした。

** : 出生数とした。

*** : $\prod p_t = p_0 \times p_1 \times p_2 \times \dots \times p_t$

① 断面的生命表の例

$$q_0 = \frac{d_0}{S_0} = \frac{2107}{612148} = 0.00344$$

また、生存率 p_0 は、1から死亡率 q_0 を引いた

$$1 - 0.00344 = 0.99656$$

と計算できる。同様にして、1歳児の死亡率 q_1 は0.00055、生存率 p_1 は0.99945と計算でき、死亡率は、0歳から、1歳、2歳と成長するにつれ、小さくなっているのがわかる。

次に、各年齢の死亡率は今後も変わらないと想定して、生まれたばかりの子が3歳ちょうどになるまで生存する確率を考える。これを生後から3歳までの累積生存率という。3歳まで生存するには、生後1年間を生存し、1歳から2歳までの1年間も生存し、2歳から3歳までの1年間も生存しなくてはならない。3期間いずれも生存する確率 P_2 は、

$$P_2 = p_0 \times p_1 \times p_2 \\ = 0.99656 \times 0.99945 \times 0.99967 = 0.99568$$

である。生き続けるということは大変である。一方、1からこの累積生存率を引いたものが、累積死亡率となる。

わが国男子の生まれてから3歳になるまでの累積死亡率（3区間のうちのいずれかで死亡する確率）は0.00432（ $= 1 - 0.99568$ ）と推定できた。同じように、ちょうど30歳からの3年間の累積生存率を知りたいければ、 $p_{30} \times p_{31} \times p_{32}$ を計算すればよい。

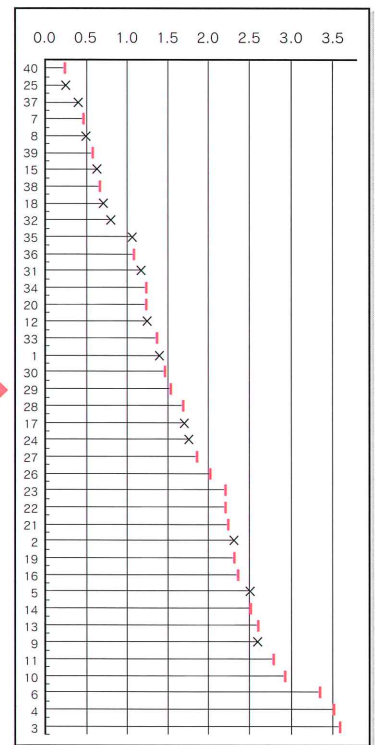
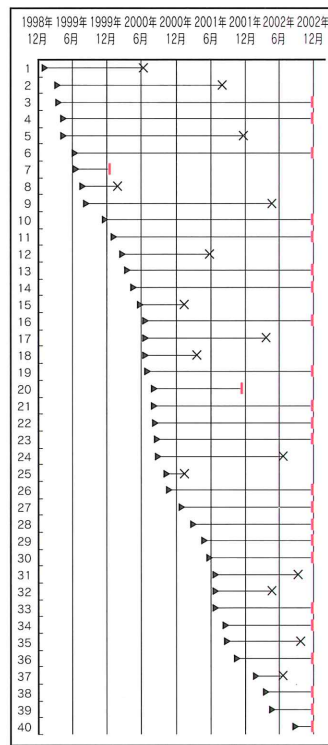
厚生労働省が発表する生命表では、その年の1月1日に t 歳ちょうどの人が1年間でのどのくらい死亡するかをできるだけ正確に推定するために、10月1日に行われる国勢調査人口を補正したり、出生や死亡の届出遅れなどの補正を行い、各死亡率を算出している²⁾。その生命表において、今後も当該年と同じ状況で死亡が起きると仮定し、 t 歳

ちよどの人が生存する期待年数を算出している。それが t 歳での平均余命であり、0歳での平均余命を平均寿命という。

ここでは、各個人は1年間という限られた観察時間でのみ観察されており、この方法で作成された生命表は共通の人々を観察した結果ではなく、各年齢階級は別々の集団から成り立っている。このような生命表を「断面的生命表」とよぶ。

経時的生命表

コホート研究や臨床試験のように、すべての個人を開始時点から追跡観察したデータから作成される生命表を、「経時的生命表」とよぶ。医学論文に登場する生命表法の多くは経時的なものである。次に経時的生命表を作成してみよう。



② 治療開始から治癒までの時間的経過

治療開始後 年数	区間当初 生存数	区間内 死亡数	区間内 打ち切り数	死亡率 (/半年)	生存率 (/半年)	累積 生存率	累積 死亡率
t	S_t	d_t	w_t	q_t	p_t	P_t	Q_t
0.0- $<$ 0.5	40	3	2	0.0769	0.9231	0.9231	0.0769
0.5- $<$ 1.0	35	3	2	0.0882	0.9118	0.8417	0.1583
1.0- $<$ 1.5	30	4	5	0.1455	0.8545	0.7192	0.2808
1.5- $<$ 2.0	21	2	3	0.1026	0.8974	0.6454	0.3546
2.0- $<$ 2.5	16	1	6	0.0769	0.9231	0.5958	0.4042
2.5- $<$ 3.0	9	1	1	0.1176	0.8824	0.5257	0.4743
3.0- $<$ 3.5	7	1	4	0.2000	0.8000	0.4206	0.5794

区間当初の生存数、区間内死亡数、区間内打ち切り数は、②(右図)から数えることができる。

③ 生命表法を用いた治療開始後の累積生存率

②の左図は、ある疾患の治療開始後の経過を示した仮想データである。縦軸が患者No.を表している。患者No.1では、1999年1月に治療開始(▶)し、2000年6月に死亡事象(×)が発生した。全40例のうち死亡事象が発生したのは15例である。No.3, No.4, No.6などは、研究全体の最終観察時点である2002年末までに死亡が発生せず、その時点で生存が確認されている。また、No.7やNo.20は、その疾患や治療とは関係のない転居や事故のために転院があり、以降の追跡観察ができなかった例である。このように、事象が発生する前に観察が中断されてしまうことを、「観察打ち切り(withdrawal, censored)」とよぶ。いずれも生存を最後に確認した時点(■)の印で示している。

次に、各例の治療開始時点から観察終了時点(死亡事象発生時点もしくは観察打ち切り時点のいずれか)までの期間を計算し、その期間の短い例から順に示したものが②の右図である。この右図をもとに、生命表を作成してみた(③)。表の左欄から治療開始後の時間的区間 t 、区間当初の生存数(区間最初に観察対象となっていた人数) S_t 、区間内での死亡発生数 d_t 、

区間内での打ち切り数 w_t を示すが、これらは②(右図)から数えることができる。各区間での半年あたりの死亡率 q_t は、区間中に発生した死亡数を該当区間の観察人数で割れば求められる。そこで区間中の観察打ち切り数の扱いが問題となる。ここでは、

$$q_t = \frac{d_t}{S_t - 0.5w_t}$$

と算出している。区間の初めで起きた打ち切りも、最後に起きた打ち切りも、区間のちょうど半分まで観察したと考え、式の分母である区間中観察人数は、区間当初の人数から、打ち切り例数の半分を引いている。ほかにも、打ち切り例を

区間の当初に起きたと考え分母からすべて引く方法や、区間最後に起きたとして分母から引かない方法もある。

各区間での生存率 p_t は $1 - q_t$ で、あとは断面的生命表と同じ手順で、累積生存率、累積死亡率を計算すればよい。この方法は、「生命表法(life table法, Cutler-Ederer法)」とよばれ、生命保険分野での生命表作成で古典的に用いられてきたので生命保険数理法(actuarial法)ともいう。

各累積生存率の標準誤差は、グリーンウッド(Greenwood)の式*から算出することができる。

カプラン・マイヤー法での生存曲線

生命表法を利用すれば、各個人の観察期間がまちまちであっても、非可逆的事象の時間的推移を検討できる。しかしながら、生命表法では時間区間の設

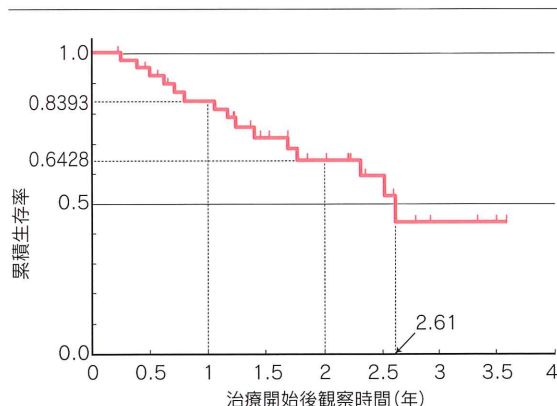
* グリーンウッドの式

$$SE(P_t) = P_t \sqrt{\sum \frac{d_t}{S_t(S_t - d_t)}}$$

定は任意であり、あまり大きな区間を設定しすぎると、事象発生の時間的特徴を見逃す可能性がある。たとえば、人の生存時間（寿命）を解析するのに1区間を100年に設定して生命表を作成しても、その生存時間の特徴をとらえることはできない。そこで、時間区間の大きさに依存しない累積生存率の算出法として、「カプラン・マイヤー法（Kaplan-Meier法、積極限推定量）」が知られている。

先ほどの②（右図）において、観察期間の短い例から順に、死亡事象発生であれば $p_i = \frac{S_i - 1}{S_i}$ 、観察打ち切りであれば $p_i = 1$ として、累積生存率（ $P_i = \prod p_i$ ）を算出していく。各例の観察時間を x 軸に、累積生存率を y 軸にプロットし、直線で結ぶ。事象が発生するたびに、その時点の累積生存率は下がり、階段状の図が描かれる。観察打ち切り例では、その時点の生存曲線上に小さなヒゲの印を立てて位置を示す。

②（右図）の仮想事例を使って、カプラン・マイヤー法で累積生存率の推移を描いたものが④である。治療開始時の累積生存率は1なので、 $x = 0, y = 1$ から階段は始まる。最も観察期間の短い患者No.40は観察打ち切り例で、治療開始0.23年後に打ち切りとなったが、その時点で $P_1 = p_1 = 40 / 40 = 1$ と累積生存率は1のままである。打ち切りの時点にヒゲだけ立てておく。その後、2番目に観察期間が短いNo.25の死亡が起き



④ カプラン・マイヤー法による累積生存率の推移図

るまでは累積生存率は1のままだが、その死亡が発生した瞬間に治療開始後0.25年時点での累積生存率は $P_2 = p_1 \times p_2 = 1 \times (38 / 39) = 0.9744$ となる。これが最初の下り階段ステップとなる。3番目に観察期間が短いNo.37の死亡が起きるまでは、その累積生存率0.9744が続き、0.39年後のNo.37の死亡発生時点で、累積生存率は $P_3 = p_1 \times p_2 \times p_3 = 1 \times 0.9744 \times (37 / 38) = 0.9487$ と下がる。このように、順次 p_i を掛けて累積生存率を算出して直線でつなぐ。

累積生存率の推定値は、治療開始1年後の時点で0.8393、2年後の時点で0.6428と、生命表法での累積生存率③とほぼ同じ値となった。また、④で累積生存率が0.5となる時点「50%生存時間（median survival time）」という。観察開始後に、全体の半数の人に事象が発生する時間を表す。この場合では、50%生存時間は、治療開始後2.61年であった。

生存曲線の比較

次に、喫煙者群と非喫煙者群との生存曲線や、治療A法群とB法群との累積治癒率の推移といった2群間の比較をすることを考えよう。各群での50%生存時間といったある1時点の指標を比べるのも一案だが、生存曲線全体の特徴を比べる方法として、ログランク検定と、一般化ウィルコクソン検定がよく用いられる^{3, 4, 5)}。いずれの検定手順も手計算でやると、観察数が大きい場合には気が遠くなる（だけでなく、たいがいは計算を間違える）。実際はHALWIN、SASといった統計パッケージプログラムを用いて解析するのが普通である。医学論文でも、統計解析方法の項では使用した統計パッケージ名が記載される。ここでは、生存時間データで汎用される検定での、手順の概要と特徴を理解してもらいたい。

◆ログラंक検定

本連載「EBNのための統計の読み方⑤交絡因子の影響を除く方法」⁶⁾において、交絡因子で層別化した四分表の検定で登場したマンテル・ヘンツェル検定を思い出してほしい**（期待度数と観察度数とを引き算したとか、 Σ とかいって全層の和を取ったとかの記憶が残っていれば…）。このときの交絡因子の代わりに時間区間を層と考え、各区間で2群での生存数・死亡数から四分表を作れば、生存時間データで同じ検定統計量が利用できる。②で生命表を作成したときのように、たとえば6か月ごとに時間区間を区切って、観察開始時から6か月までを第1区間、次の6か月を第2区間、……と層を作ればよい。連載⑤の交絡因子のときと記号をそろえ、第k区間での四分表を⑤とするとマンテル・ヘンツェル検定統計量 χ^2_{CMH} は、

$$\chi^2_{CMH} = \frac{\left(\left| \sum_k \left(a_k - \frac{x_k m_k}{N_k} \right) \right| - 0.5 \right)^2}{\sum_k \frac{m_k n_k x_k y_k}{N_k^2 (N_k - 1)}}$$

となり、この値は自由度1の χ^2 分布に従う。この値

が3.84（有意水準5%）より大きければ、2群の生存曲線に有意な違いがあるとする。

第k区間内での観察打ち切り例は、区間当初で打ち切られたとして第k区間の四分表に含めない場合と、第k区間終了直前まで観察したとして生存に入れる場合とが考えられる。また、6か月といった特定の時間的区間で区切るのではなく、いずれかの群で死亡例が発生するたびに区間を区切って、その時点での死亡数と生存数で四分表を作り、マンテル・ヘンツェル検定統計量を算出すると、「コックス・マンテル検定（Cox-Mantel test）」となる。時間区間が細かくて特に分母の分散の計算が大変なので、近似的な式として、

$$\chi^2_L = \frac{\left\{ \sum_k \left(a_k - \frac{x_k m_k}{N_k} \right) \right\}^2}{\sum_k \frac{x_k m_k}{N_k}} + \frac{\left\{ \sum_k \left(b_k - \frac{y_k n_k}{N_k} \right) \right\}^2}{\sum_k \frac{y_k n_k}{N_k}}$$

の式がある。各群の観察度数と期待度数の差の2乗を期待度数で割った値の和となっているので、計算がずいぶん楽である。これを、「ログラंक検定（log-rank test）」という。この値も自由度1の χ^2 分布で、3.84（有意水準5%）より大きければ2群の生存曲線に有意な違いがあるとする。コックス・マンテル検定と比べて、若干は保守的になる（有意になりにくい）。

◆一般化ウィルコクソン検定

「一般化ウィルコクソン検定（generalized Wilcoxon test）」は、順序分類データを2群間で比較するウィルコクソン順位和検定***（Wilcoxon rank sum test）を、生存時間データに一般化した

	第k区間	死亡数	生存数	観察人数
第1区間	A群	a_k	b_k	m_k
	B群	c_k	d_k	n_k
	計	x_k	y_k	N_k

⑤ 第k区間の四分表

** 一般に、性別、年齢、喫煙習慣などの交絡因子の影響を除き、評価したい要因の影響を正しく推定するための方法に、層別解析という方法がある。これは、交絡因子の各カテゴリー（層）ごとに集計を行い、各層内では交絡因子が均一になるようにすることで、交絡因子の影響を除くという方法である。……このような場合、複数の四分表をまとめて、各層に共通するオッズ比（共通オッズ比）を推定するための近似的な方法が、Mantel-Haenszelによって提唱された。

*** ウィルコクソン順位和検定

データを大きさの順に並べて、大小関係としての順位データとして分析する方法。順序分類尺度で測ったデータはもちろん、計量値データでもはずれ値がある場合など広く利用される。

対応のない2群を比較する検定法であり、両群の例数が異なる場合に拡張されたものを、Mann-WhitneyのU検定とよぶこともある。

ものである。先ほどの治療A法群とB法群を例にすると、A群の全例 (n_a 人) とB群の全例 (n_b 人) との間を総当たりで、各個人の生存時間の長短を比較していく。何年長いといった生存時間の差の大きさは問題にせず、長いか短いかの大小関係のみに注目する。A群の患者A₁の生存時間を、B群の患者B₁の生存時間と比べて、長ければ1点、同じならば0点、短ければ-1点とする。観察打ち切りなどで長短判断できない場合は0点とする。

たとえば、生存時間が2年の観察打ち切りの患者A₁は、生存時間1年で死亡した患者B₁との比較では明らかに長いと判断でき1点だが、生存時間3年で死亡した患者B₂との比較では長短が判断できず0点となる。同様に、患者B₃, B₄, B₅, …, B_{n_b}の生存時間も比べて、それぞれ1点, 0点, もしくは-1点とし、患者A₁でのB群全患者B₁, …, B_{n_b}との総当たり比較の合計点を w_1 とする。もし、患者A₁の生存時間がB群のすべての患者より長ければ w_1 は n_b 点、すべての患者より短ければ w_1 は $-n_b$ 点となる。同じようにして、患者A₂でのB群全患者との総当たり比較の合計点を w_2 とする。こうして w_{n_a} までを出して、A群全体の群合計点 $w = w_1 + w_2 + \dots + w_{n_a}$ を算出する。B群側から総当たりを行っても、合計点 w の+-の符号は逆だが、その絶対値は同じとなる。

次に、群には関係なく、($n_a + n_b$) の全例で総当たりをして得点を求める。患者A₁について、患者A₂, …, A_{n_a}, B₁, …, B_{n_b}との比較の合計点を W_1 とする。同じようにして患者A₂の W_2 から、患者B_{n_b}の W_{n_b} まで求める。それらの2乗の合計 $\sum W_i^2$ を計算する。そこで、一般化ウィルコクソン検定の検定統計量 z は、

$$z = \frac{|w|}{\sqrt{\frac{n_a n_b \sum W_i^2}{(n_a + n_b)(n_a + n_b - 1)}}$$

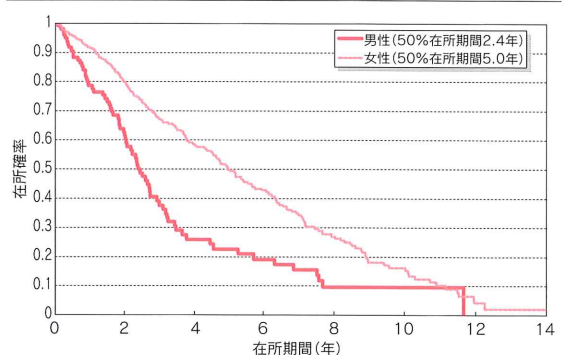
となり、標準正規分布に従う。この値が1.960 (有意水準両側5%) より大きければ、2群の生存曲線には統計学的に有意な違いがあるとする。

◆ 2つの検定の特徴

一見すると、ログランク検定と一般化ウィルコクソン検定は、まったく異なる検定法に見える。しかし、時間区間への重み付けの考え方を導入すると、両者は統一した表現形式がとれることがわかっている。ログランク検定では、どの時間区間も重みは等しい。一方、一般化ウィルコクソン検定では観察対象数に比例した重みとなるので、観察開始直後の区間で最も重みが大きく、時間経過とともに重みは小さくなる。このことから、ログランク検定は生存曲線の後期での差を、一般化ウィルコクソン検定では生存曲線の初期での差を検出するのに感度がよいといわれる^{4,5)}。

⑥の生存曲線は、ある特別養護老人ホームにおいて、入所から退所までの在所期間を調べた事例である⁷⁾。入所時年齢が80歳代の高齢者430人 (男性96人, 女性334人) の在所の様子を、男女別に Kaplan-Meier 法での生存曲線として表している。図が煩雑になるため観察打ち切りの印はつけていない。50%在所期間は男性で2.4年、女性で5.0年と、女性のほうが在所期間は長そうである。

前述の検定法を用いて2つの生存曲線全体を比較してみると、ログランク検定では $\chi^2 = 21.09$, $p < 0.001$, また一般化ウィルコクソン検定では $z = 5.221$,



⑥ 入所時80歳代の高齢者における在所期間 (男女別)

$p < 0.001$ と、いずれも有意となり、男性と女性の在所期間は明らかに異なっているといえる。

おわりに

⑥の例のように、生存曲線を推定し、検定によって比較すると、性別という一つの要因についての生存時間の比較ができる。しかし、生存時間に影響を与える数多くの要因（いわゆる予後要因）を一様に

検討する、もしくはほかの要因で調整して比較するためには、比例ハザード・モデルなどモデルを用いた解析が行われる。いずれのモデルでも生存時間の分布やハザードについていくつか条件が必要となり、また最尤法を用いた少々複雑な反復計算なども必要となる。手法の詳細は他書に譲るが^{3,4,5)}、多くの統計パッケージプログラムに、これらモデルを用いた解析法が用意されているので利用いただきたい。

文献

- 1) マックス・アラン・コリンズ著、小田川佳子訳：ターク・エンジェル最終戦争、角川文庫、2003。
- 2) 厚生労働省大臣官房統計情報部編：第19回生命表。厚生統計協会；2002。
- 3) 柳井晴夫、高木廣文編：新版看護学全書 統計学。メチカルフレンド社；1995。
- 4) 大橋靖雄、浜田知久馬：生存時間解析。東大出版会；1995。
- 5) 佐久間昭：医学統計Q & A。金原出版；1985。
- 6) 高木廣文、林邦彦：交絡因子の影響を除く方法。EB NURSING 2002；2（1）：94-100。
- 7) 痴呆性高齢者の予後追跡調査委員会：痴呆性高齢者の長期介護に関する研究のうち痴呆性高齢者の予後追跡調査研究報告書ー平成14年度老人保健健康増進事業等事業報告書。浴風会高齢者痴呆介護研究・研修東京センター、平成15年3月。

INFORMATION

(財)ライフ・プランニング・センター 国際フォーラム

日時：2004年8月28日(土)13:00～18:00、
29日(日)9:00～16:30

会場：聖路加看護大学ホール（東京都中央区明石町1-10）

テーマ：ナースによるフィジカルアセスメントの実践

プログラム：【28日】開会の言葉（道場信孝）／「日米の看護教育の比較に触れて」（日野原重明）／「米国におけるナースプラクティショナー発展の歴史—過去40年間の経緯と展望」（Patrice Nicholas）／「米国におけるナースプラクティショナーの役割：看護における高いレベルのプラクティスとしての役割に関する展望」（Margery Chisholm）／「UCSF看護学部におけるナースプラクティショナー教育と日本での展望」（操華子）／「日本の看護教育におけるフィジカルアセスメントに関する教育の現状と問題点」（山内豊明）／「チームケアの中で求められる看護師のアセスメント能力—在宅医療の現場から」（和田忠志）【29日】「米国におけるナースのための高度なアセスメントの教育法：問診、身体所見、鑑別診断、

マネージメント」（P. Nicholas）／「ナースプラクティショナー養成カリキュラム—その専門分野と役割」（M. Chisholm, P. Nicholas）／「高いレベルのプラクティス技術の修得をどのように展開するか—LPCでの試み」（道場信孝、石清水由紀子）／総括討論／記念講演「オスラーに見る看護師像—いかにそれを具現するか」（日野原重明）
*同時通訳サービスあり

参加費：LPC会員7,350円、非会員10,500円
申込方法：TELまたはFAXで参加希望を連絡し定員以内かを確認の上、郵便振込みにて参加費をお支払いください。郵便振込み取り扱ひ表の依頼人欄に、郵便番号・住所・氏名・電話番号を、通信欄に、国際フォーラム参加希望・職業・所属施設名・（LPC会員の方は）会員番号ご記入ください。ご入金いただいた参加費は返金できません。

郵便振込先：口座番号：00150-7-185081、加入者名：(財)ライフ・プランニング・センター

定員：350名（定員になり次第締め切り）

問い合わせ先：〒102-0093 東京都千代田区平河町2-7-5 砂防会館5階 (財)ライフ・プランニング・センター

TEL：03-3265-1907、FAX：03-3265-1909