

信頼性工学 最終課題

中央大学理工学部ビジネスデータサイエンス学科

23D7104001I 高木悠人

作成日: 2025/07/15

1. 解析の目的

本研究の目的は、薬物依存症患者に対する各種治療介入が再使用（リラプス）確率に及ぼす影響を定量的に評価し、より効率的かつ効果的な治療法の開発に資する知見を得ることだ。近年、フェンタニルをはじめとする高い中毒性を有する合成オピオイドの乱用が世界規模で急増し、従来の規制対象に含まれない合成カンナビノイド（CBN 等）が若年層を中心に蔓延している現状がある。こうした新興薬物の流行は、公衆衛生上の重大な課題を提起しており、再依存リスクの抑制は社会的要請となっている。

一方で、依存症治療には薬物療法、認知行動療法、集団カウンセリングなど多様なアプローチが存在し、その組み合わせや投与タイミングによって再使用防止効果に大きな差が生じる可能性がある。しかし、臨床現場では標準化された評価指標が乏しく、各治療法の相対的有効性を比較検討したエビデンスはまだ限定的であると考ええる。

そこで本課題では、国内外のリハビリ施設や保健センターから収集した追跡データを用い、生存時間解析（Cox 回帰モデル等）によって治療開始後の「再使用までの時間」と「再使用発生の有無」を解析する。具体的には、治療群ごとのハザード比を推定し、年齢、人種、ヘロイン、コカイン使用歴などの共変量で調整した上で再使用リスクの有意差を検証する。その結果を基に、再依存率を最も低減させる治療プロトコルの特徴を抽出し、今後の臨床ガイドライン整備や公共政策立案に貢献することを目指す。

2. 解析方法

本研究では、薬物依存症患者の治療介入効果を再使用リスクの観点から評価するため、以下の手順で統計解析を実施した。

(ア) データ前処理

収集した追跡データから欠測値を含む症例を除外し、主要解析変数（再使用までの時間 `time`、再使用イベント有無 `event`、治療群 `indicator`、年齢・人種などの共変量）を抽出した。カテゴリ変数は Python の `category` 型に変換し、連続変数は必要に応じて中央値カットを行った。そのうえで、`ydata_profiling` を用いた探索的データ分析(EDA)も実施した。

(イ) 生存曲線の描画

まず Kaplan–Meier 法により全体および各治療群の生存曲線（再使用までの未発生確率曲線）を算出し、ログランク検定、グローバル検定、ペアワイズ検定によって群間比較を行った。

(ウ) 単変量 Cox 回帰解析

各治療変数および主要共変量について単変量 Cox 比例ハザードモデルを当てはめ、Wald 検定による係数の有意性と Schoenfeld 残差検定による比例ハザード性の妥当性を確認した。

(エ) 多変量 Cox 回帰解析

単変量解析で有意かつ比例ハザード性を満たした変数を選択し、年齢・人種等を共調整した多変量 Cox モデルを構築した。モデル全体の適合度は尤度比検定（LRT）で評価し、各係数の推定値はハザード比（ $\exp(\text{coef})$ ）と 95%信頼区間で報告した。

(オ) 比例ハザード性の検定

多変量モデルでは `proportional_hazard_test()` を用いて Schoenfeld 残差検定を再度実施し、時間依存性がないことを確認した。

(カ) ソフトウェア

解析は主に Python（pandas, lifelines, matplotlib）上で行い、コードと設定を GitHub 上でバージョン管理した。Python のバージョンは 3.11.7 を使用し pipenv での仮想環境構築を用いた。そして、本解析で利用したパッケージは requirements.txt にまとめた。

3. 解析結果

(ア) 探索的データ分析

本データの概要把握として、探索的データ分析を行った。

1. テーブル形式

まず、データテーブルの列情報について以下に示す。

表 1. データテーブルの各列情報と形式

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 628 entries, 0 to 627
Data columns (total 11 columns):
#   Column      Non-Null Count  Dtype
---  ---
0   pid         628 non-null    int64
1   event       628 non-null    int64
2   time        628 non-null    int64
3   fac_heroic  628 non-null    category
4   fac_ivhx    628 non-null    category
5   fac_race    628 non-null    category
6   fac_treat   628 non-null    category
7   fac_site    628 non-null    category
8   num_age     623 non-null    float64
9   num_beck    595 non-null    float64
10  num_ndrugtx 611 non-null    float64
dtypes: category(5), float64(3), int64(3)
memory usage: 32.8 KB
```

表 2. データテーブル先頭 3 行

	Unnamed: 0	event	time	fac_heroic	fac_ivhx	fac_race	fac_treat	fac_site	num_age	num_beck	num_ndrugtx
0	0	1	188	4.0	3.0	0.0	1	0	39.0	9.0	1.0
1	1	1	26	4.0	2.0	0.0	1	0	33.0	34.0	8.0
2	2	1	207	2.0	3.0	0.0	1	0	33.0	10.0	3.0

上記をもとに、各列情報を以下にまとめる。

(1) Unnamed

列,レコード番号で csv 形式として保存する際に追加した。

(2) event

イベント発生における bool 値で、event=1 の場合がイベント発生、event=0 で観察期間内のイベント未発生と観察中断を示す。本データについてのイベントとは、薬物中毒に対する治療後に再度使用することである。

(3) time

event 発生または観察中止、中断までの時間である。具体的には、event=1 におけるレコードの time は、治療後の再使用までの時間を意味する。

(4) fac_heroic

直近 3 ヶ月のヘロイン,コカイン使用歴を示すカテゴリ変数。具体的には、以下のとおりである。

- 1: ヘロイン, コカインどちらも使用
- 2: ヘロインのみ使用
- 3: コカインのみ使用
- 4: いずれも使用せず

(5) fac_ivhx

違法薬物を静脈注射として使った経験を示すカテゴリ変数。具体的には以下の通りである。

- 1: 違法薬物における注射経験無し
- 2: 過去に注射経験はあるが、直近ではない
- 3: 直近に使用履歴がある

(6) fac_race

人種を示すカテゴリ変数で、具体的には以下のとおりである。

- 0: 白人
- 1: その他

(7) fac_treat

治療プログラム郡を示すカテゴリ変数である。治療プログラムについては、ランダムな割り付けを行っている。具体的な治療プログラムは以下のとおりである。

- 0: 短期プログラム
- 1: 長期プログラム

(8) fac_site

治療施設を示すカテゴリ変数である。本データは 2 つの施設での治療に注目しており、具体的な対応は以下の通りである。

- 0: 施設 A
- 1: 施設 B

(9) num_age

年齢を示す数値列である。

(10) num_beck

自己記入式の抑うつ尺度で過去 1 週間の抑うつ症状の強さを定量的に評価できる。スコアに対する抑うつ状態は以下のとおりである。

表 3. 抑うつ尺度

スコア	状態
0~9	正常~軽度抑うつ
10~18	軽度~中程度抑うつ
19~29	中程度以上抑うつ
30 以上	重度抑うつ

(11) num_ndrugtx

薬物治療プログラムの過去受講回数を示す数値列。そのため、受講回数が多いほど、中毒性が高く難治度、慢性度が高いと推測できる。

2. 欠損値について

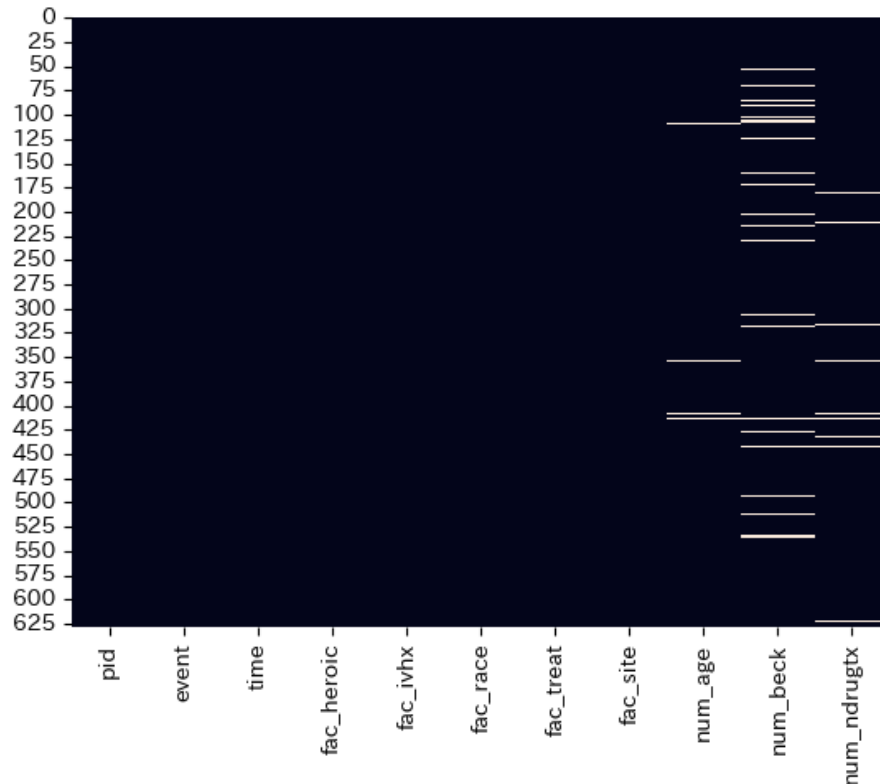


図 1. 欠損値について

上記の図を見ると、「 num_beck 」で多くの欠損値が存在していることがわかる。ただし、レコード数が 628 であり欠損レコードは 33 であり、欠損値部分が大きな影響を与えるとは考えにくい。また、うつ状態であったかを示す列であるため、本研究の際にすべての被験者に調査することが難しかったのも想定できる。したがって、欠損値をもつレコードについては除外して解析を進めることとした。

3. 再使用, 未使用割合

以下の図2は、調査対象者における再使用,未使用および観測打ち切り者割合を示した円グラフである。ただし、未再使用とは観測基準期間(約500日後)超過による打ち切り対象者であるため、観測基準期間内に約8割の麻薬中毒者が再度使用に至るといえるとともに、観測基準期間以降に使用した場合も未再使用に含まれる。

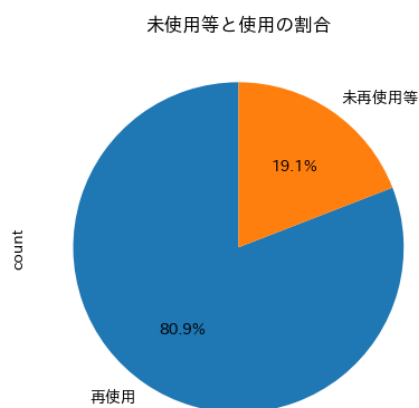


図2. 再使用, 未使用等割合の推移

次に、打ち切り発生タイミングについて観察する。図3に、観察期間におけるイベント発生について示す。図3を見ると、500日以降にしか打ち切りが発生しておらず、観察不可による打ち切りではなく意図的な観察終了による打ち切りであると想定できる。

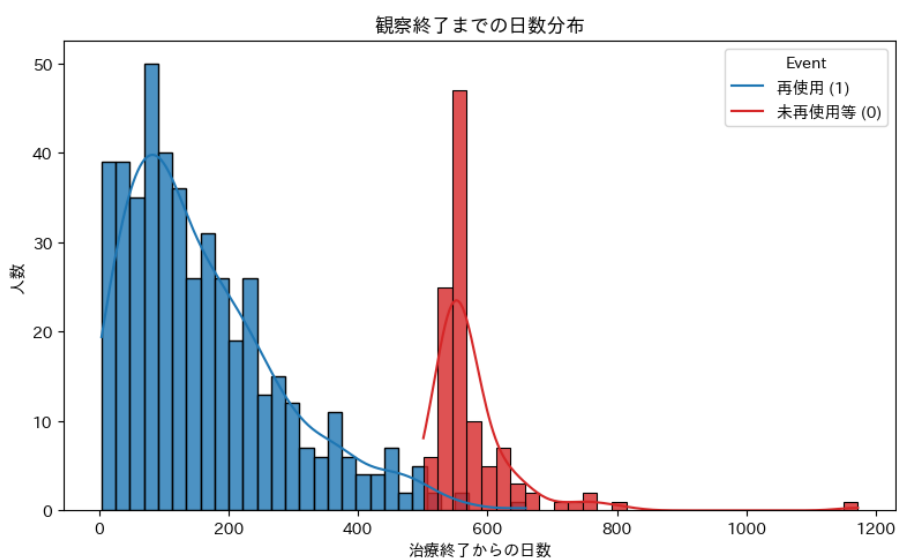


図3. 再使用までの時間分布

4. 各変数同士の相関について

次に、各変数同士の相関について観察する。共変量の相関行列について、図4に示す。

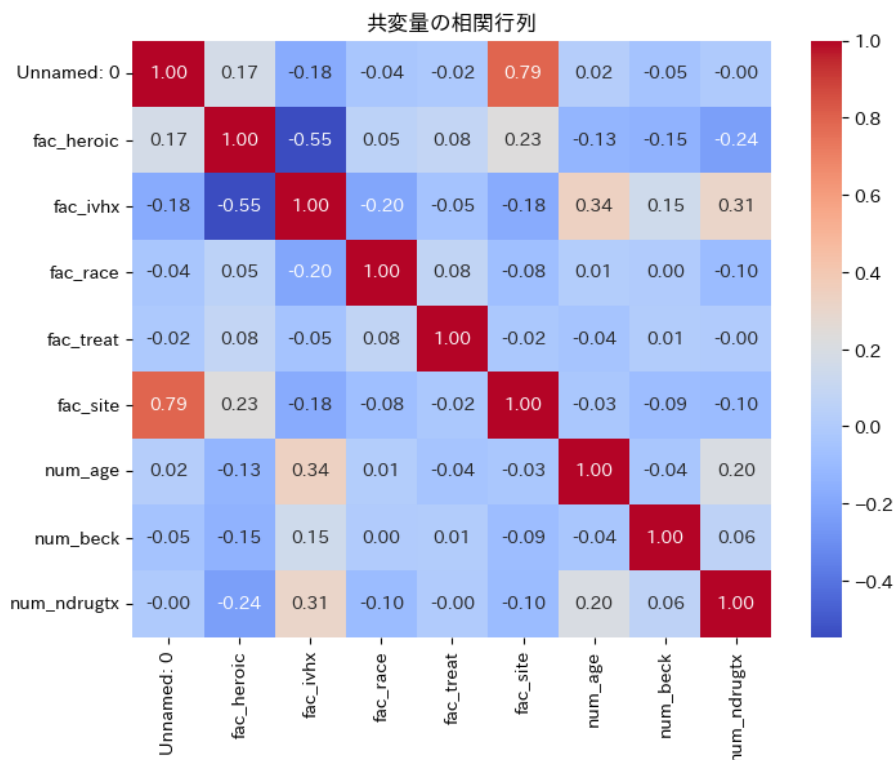


図4. 共変量の相関行列

上記の図を見ると、「fac_ivhx(静脈注射使用歴)」と「fac_heroinc(3ヶ月以内のヘロインコカイン使用歴)」に負の相関があることや、「num_age(年齢)」と「fac_ivhx(静脈注射使用歴)」に正の相関があることがわかる。静脈注射使用歴については、カテゴリ変数であり数値が上がるごとに直近の使用が多いといえる。一方で、3ヶ月以内のヘロインコカイン使用歴についてはカテゴリ変数であり数値が上がるごとに、使用種類が少なくなる。具体的には、1だとヘロイン、コカイン両方使用であるが、4(最大)だといずれも使用していないという扱いである。つまり、静脈注射を使用している薬物中毒患者が直近3ヶ月でヘロイン、コカインを使用している傾向にあることは、理解しやすいといえる。そして、年齢と静脈注射歴の正相関についても年齢を重ねることで薬物使用回数(静脈注射回数)も増すと想像できる。ただし、そのような傾向についての有意な関係、再使用までの期間への有意と言える影響がどの程度あるのかについては調査する必要があるといえる。

上記の探索的データ分析をもとに、分析を進めることとした。

(イ) Kaplan-Meier 曲線とログランク検定

まず、それぞれの説明変数について、Time-To-Event に対して影響があるのか、可視化するため Kaplan-Meier 曲線を確認する。その際に、複数の集合ごとの曲線同士が交わらない場合は、比例ハザード性があると推察でき、時間によって、その影響が変化しないと評価できる。よって、Cox 比例ハザードモデルを利用した説明変数ごとの寄与度を評価することができるようになる。

1. fac_heroic(ヘロイン, コカインの使用歴)

まず、ヘロイン, コカイン使用歴についての Kaplan-Meier 曲線を図 5 に示す。

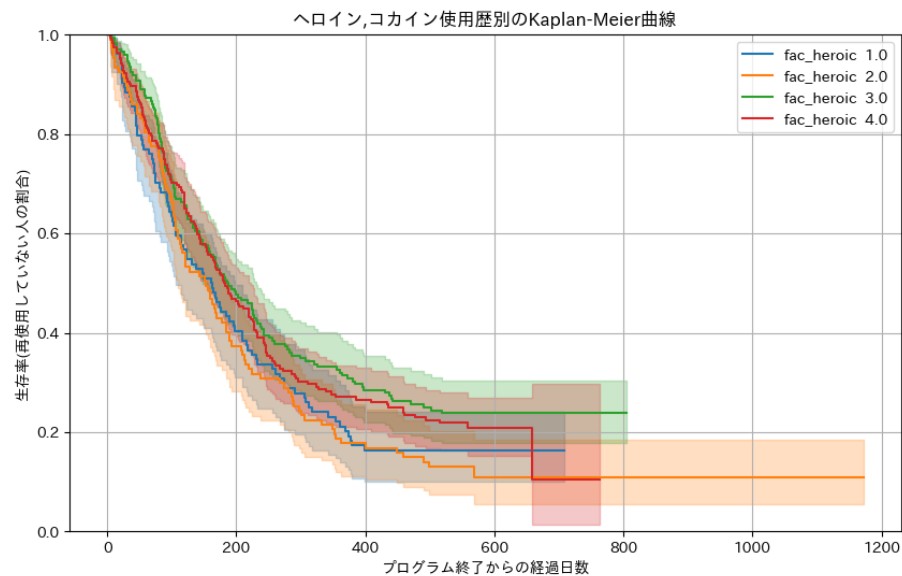


図 5. ヘロイン,コカイン使用歴による Kaplan-Meier 曲線

上記の図を見ると、1.0, 2.0 のグループ(どちらも使用とヘロインのみを利用)のほうが、生存率(期間内に再使用しない確率)が低いと推察できる。ただし、1.0 と 2.0, 3.0 と 4.0 で、曲線の推移が近い。また、David Nutt 氏らの調査による薬物の「身体的被害」「依存」「社会的被害」での評価である表 4 を見ると、ヘロインはコカインよりも、依存スコアが 0.41 ポイント高く最高スコアとなっている。よって、どちらの使用もないグループとコカインのみ使用したグループを同一とし、どちらも使用しているグループとヘロインのみを使用しているグループを同一とする上で、Kaplan-Meier 曲線を作図することとした。その結果を図 6 に示す。

表 4. 薬物の 3 観点での評価

	Physical harm				Dependence				Social harm			
	Mean	Acute	Chronic	Intravenous	Mean	Pleasure	Psychological dependence	Physical dependence	Mean	Intoxication	Social harm	Health-care costs
Heroin	2.78	2.8	2.5	3.0	3.00	3.0	3.0	3.0	2.54	1.6	3.0	3.0
Cocaine	2.33	2.0	2.0	3.0	2.39	3.0	2.8	1.3	2.17	1.8	2.5	2.3
Barbiturates	2.23	2.3	1.9	2.5	2.01	2.0	2.2	1.8	2.00	2.4	1.9	1.7
Street methadone	1.86	2.5	1.7	1.4	2.08	1.8	2.3	2.3	1.87	1.6	1.9	2.0
Alcohol	1.40	1.9	2.4	NA	1.93	2.3	1.9	1.6	2.21	2.2	2.4	2.1
Ketamine	2.00	2.1	1.7	2.1	1.54	1.9	1.7	1.0	1.69	2.0	1.5	1.5
Benzodiazepines	1.63	1.5	1.7	1.8	1.83	1.7	2.1	1.8	1.65	2.0	1.5	1.5
Amphetamine	1.81	1.3	1.8	2.4	1.67	2.0	1.9	1.1	1.50	1.4	1.5	1.6
Tobacco	1.24	0.9	2.9	0	2.21	2.3	2.6	1.8	1.42	0.8	1.1	2.4
Buprenorphine	1.60	1.2	1.3	2.3	1.64	2.0	1.5	1.5	1.49	1.6	1.5	1.4
Cannabis	0.99	0.9	2.1	0	1.51	1.9	1.7	0.8	1.50	1.7	1.3	1.5
Solvents	1.28	2.1	1.7	0	1.01	1.7	1.2	0.1	1.52	1.9	1.5	1.2
4-MTA	1.44	2.2	2.1	0	1.30	1.0	1.7	0.8	1.06	1.2	1.0	1.0
LSD	1.13	1.7	1.4	0.3	1.23	2.2	1.1	0.3	1.32	1.6	1.3	1.1
Methylphenidate	1.32	1.2	1.3	1.6	1.25	1.4	1.3	1.0	0.97	1.1	0.8	1.1
Anabolic steroids	1.45	0.8	2.0	1.7	0.88	1.1	0.8	0.8	1.13	1.3	0.8	1.3
GHB	0.86	1.4	1.2	0	1.19	1.4	1.1	1.1	1.30	1.4	1.3	1.2
Ecstasy	1.05	1.6	1.6	0	1.13	1.5	1.2	0.7	1.09	1.2	1.0	1.1
Alkyl nitrites	0.93	1.6	0.9	0.3	0.87	1.6	0.7	0.3	0.97	0.8	0.7	1.4
Khat	0.50	0.3	1.2	0	1.04	1.6	1.2	0.3	0.85	0.7	1.1	0.8

Table 3: Mean independent group scores in each of the three categories of harm, for 20 substances, ranked by their overall score, and mean scores for each of the three subscales

※参照元は参考文献に示すこととする。

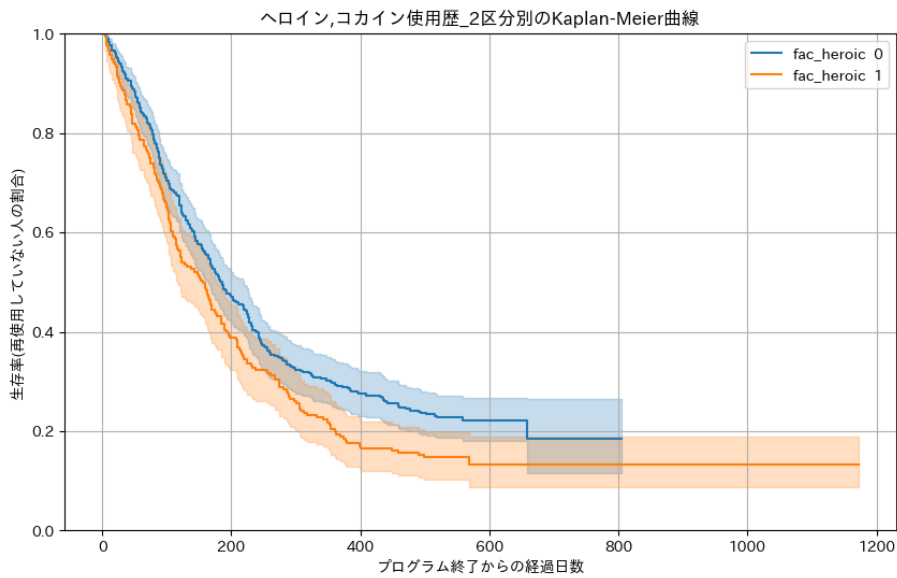


図 6. グルーピング後の Kaplan-Meier 曲線

上記を見ると、曲線同士が交差していないため比例ハザード性を保持していると推察できる。そのため、より統計的に評価するため、ログランク検定を実施する。実施したところ、以下の結果が得られた。

表 5. ログランク検定結果

	test_statistic	p	-log2(p)
0	7.036792	0.007985	6.968457

上記の検定を見ると、p 値が 0.007 となっており今回の基準である 5%を下回っている。そのため、2 群間に有意な差があると評価できる。よって、イベントに対して、有意な影響を与えている可能性が高いといえる。

2. fac_ivhx(静脈注射歴)

次に、静脈注射歴について分析する。まず、以下に Kaplan-Meier 曲線を示す。

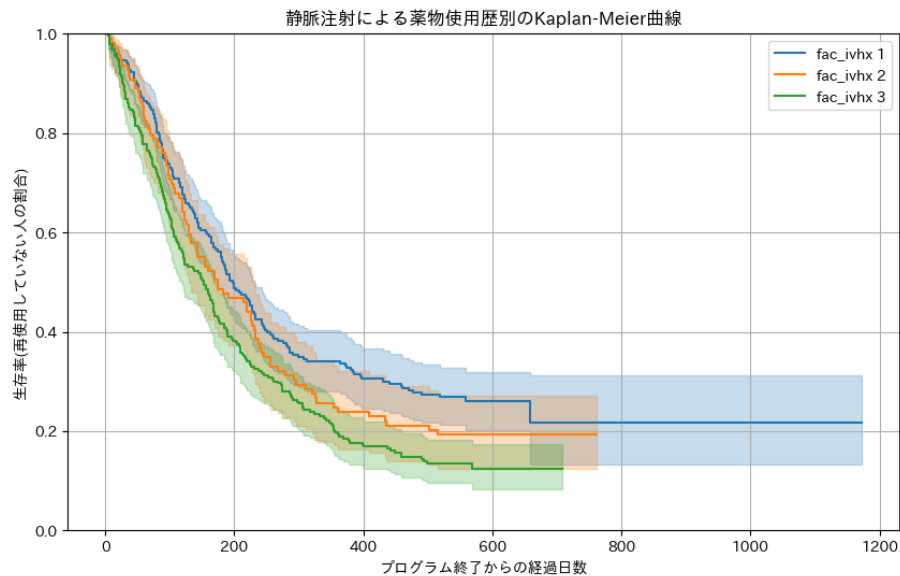


図 7. 静脈注射による薬物使用歴別の Kaplan-Meier 曲線

上記のグラフを見ると、1,2,3 とカテゴリ値が下がるにつれて生存曲線が低く推移しており重なっていないことがわかる。要するに、静脈注射の直近の使用がある方が、再使用確率が高いと言えるとともに 3 群間に比例ハザード性があると推測できる。そのため、次に 3 群に対してグローバル検定を実施することでそれら同士に有意な差があるか統計的に判断する。グローバル検定での帰無仮説、対立仮説は以下の通りである。また、検定結果を表 6 に示す。

[グローバル検定]

H_0 : すべての群の平均値は等しい

H_1 : 少なくとも 1 つの群の平均値は異なる

表 6. グローバル検定結果

	test_statistic	p	-log2(p)
0	13.723916	0.001047	9.899713

上記の結果より、有意水準 5%を下回っていることから帰無仮説を棄却し対立仮説を採択する。よって、3 群のうち 1 つでも分布の異なる群が存在すると言える。しかし、すべてが異なっているとは言えないため、ペアワイズ検定による総当たり形式のログランク検定を実施することで、どの分布間で差が生じているのか評価する。ペアワイズ検定の結果を表 7 に示す。

表 7. ペアワイズ検定結果

	3	2	1
3	NaN	0.327842	0.000681
2	0.327842	NaN	0.535419
1	0.000681	0.535419	NaN

上記を見ると、群 1 と群 3 間では有意な差があると言って良いが 1 と 2, 2 と 3 では有意水準 5%を上回るため有意な差であると評価できない。

これらより、Kaplan-Meier 曲線とグローバル検定では、差が生じていると言えたがペアワイズ検定からは補強する結果が得られなかった。そこで、p 値の大きい 1 と 2 の群をあわせ、直近使用履歴の有無の 2 カテゴリで分類し、再度 Kaplan-Meier 曲線を図示しログランク検定で評価することとする。それらの結果を図 8, 表 8 に示す。

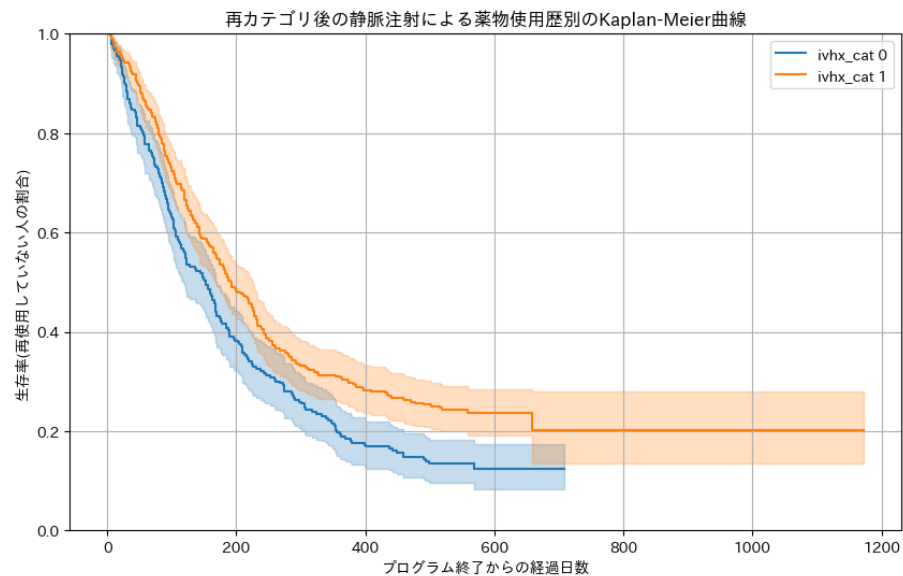


図 8. 再カテゴリ後の静脈注射による薬物使用歴別 Kaplan-Meier 曲線

表 8. ログランク検定結果

	test_statistic	p	-log2(p)
0	12.102677	0.000503	10.955735

Kaplan-Meier 曲線を見ると 2 群に差があるように見えるとともに、ログランク検定結果からも p 値が有意水準 5%を下回っているため有意な差があると評価できる。

3. fac_race(人種)

人種によるカテゴリ変数について、Kaplan-Meier 曲線を以下に示す。以下の図を見ると、推移に差があり交わりがないことがわかる。よって、白人である方が、非再使用率が低く推移している。そのため、有意な差と言えるかログランク検定で確認する。結果を表 9 に示す。表 9 を見ると p 値が 0.006 で有意水準 5%を下回っている。そのため、二群に有意差があると評価できる。

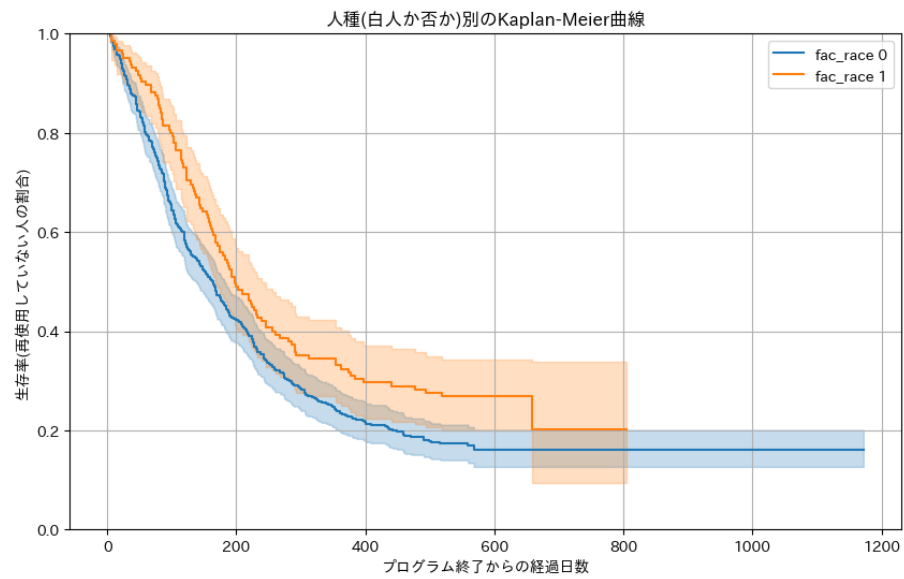


図 9. 人種(白人か否か)別の Kaplan-Meier 曲線

表 9. ログランク検定結果

	test_statistic	p	-log2(p)
0	7.431122	0.006411	7.285339

4. fac_treat(治療プログラムの種類)

治療プログラムの期間について、生存関数の推移に影響を与えているか確認するため、Kaplan-Meier 曲線を図 10 に示す。

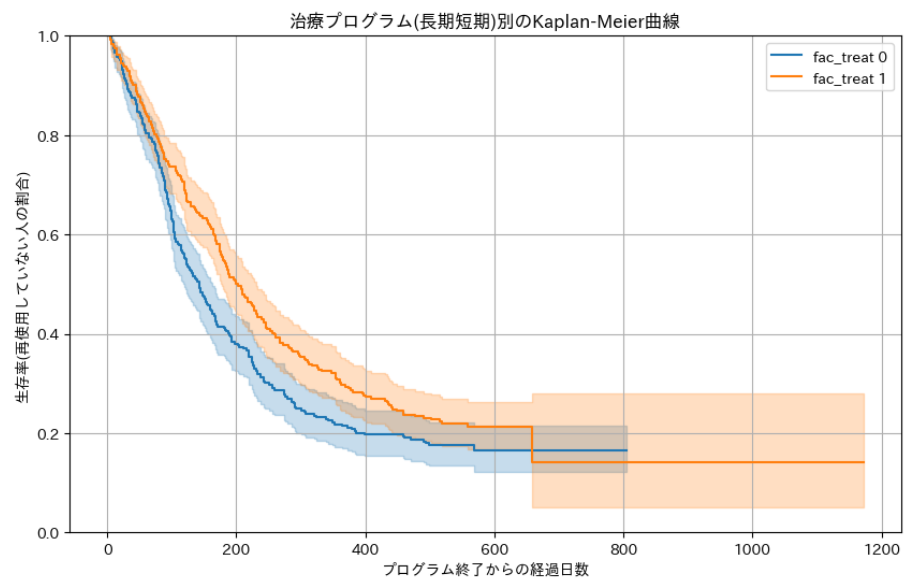


図 10. 治療プログラム期間別の Kaplan-Meier 曲線

上記を見ると、fac_treat:0 すなわち、短期プログラムの生存関数の方が低く推移していることがわかる。よって、長期期間プログラムに参加することで、更生につながっていると推測できる。ただし、期間別に関して有意な差があるのかについて確認する必要があるため、ログランク検定を実施する。実施結果を表 10 に示す。

表 10. ログランク検定結果

	test_statistic	p	-log2(p)
0	6.528395	0.010617	6.557539

表 10 を見ると、p 値が 0.01 で有意水準 5%を下回っている。よって、帰無仮説を棄却し対立仮説を採択する。よって、2 群間に有意な差があるといえる。

5. fac_site(治療施設)

次に、治療施設の違いによる生存関数推移の変化を Kaplan-Meier 曲線で確認する。曲線を図 11 に示す。

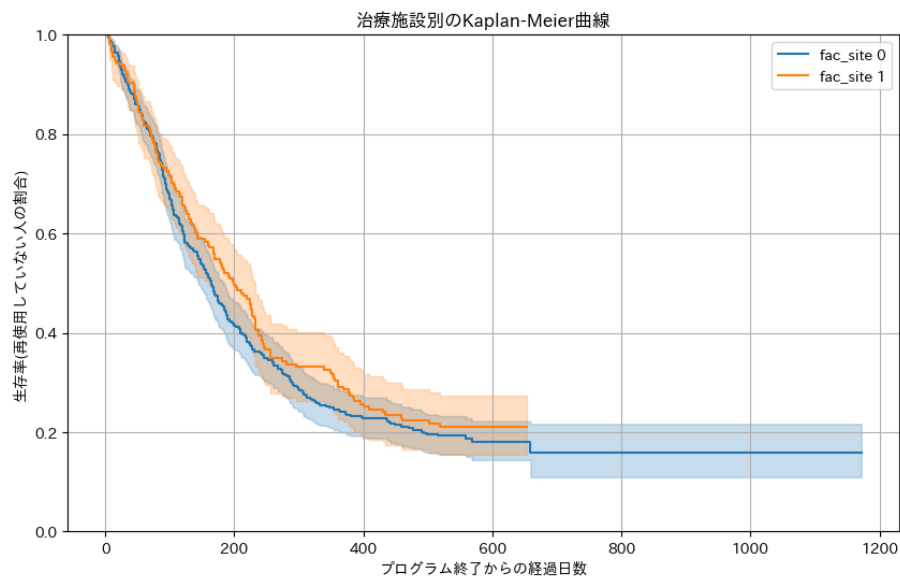


図 11. 治療施設別の Kaplan-Meier 曲線

上記の図を見ると、施設によって曲線の推移に差があるように視認できる。ただし、250 日付近で接近している点等もあり、差が有意であると断言できない。そのため、ログランク検定を用いて有意な差なのか確認す

る。その結果を表 11 に示す。

表 11. ログランク検定結果

	test_statistic	p	-log2(p)
0	1.091169	0.296212	1.755296

表 11 を見ると、p 値が 0.29 であり有意水準 5%を上回っている。したがって、KM 曲線で確認できた差が有意と断言することはできない。

6. age_cat(年齢区分)

次に、年齢による影響を可視化する。年齢は数値データでありカテゴリカルデータではない。そのため、中央値を用いて区分を定めてカテゴリカルデータに変換し、2 区分に対して Kaplan-Meier 曲線を図 12 にプロットした。

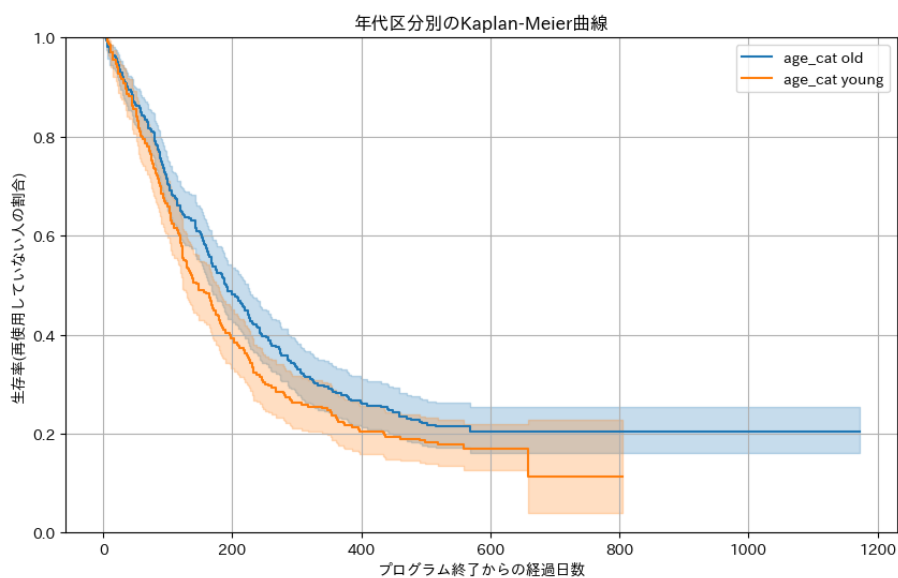


図 12. 年代区分別の Kaplan-Meier 曲線

上記を見ると、年齢が中央値(32.0)以上の場合のほうが、再使用確率が低く推移していることがわかる。また、推移が交差していないため、時間に依存せず比例ハザード性があると推測できる。ここで、2 群に有意差があるかログランク検定を実施する。検定結果を表 12 に示す。

表 12. ログランク検定結果

	test_statistic	p	-log2(p)
0	3.90029	0.048278	4.372497

上記の表を見ると、p 値が有意水準 5%を下回っている。よって、帰無仮説を棄却し対立仮説を採択する。よって、2 群に有意な差があると評価できた。

7. beck_cat(抑うつスコアレベル)

次に、抑うつスコアレベルでの Kaplan-Meier 曲線を作成する。元データでは、抑うつスコアになっているが、KM 曲線を描くために表 3 のスコアレベルで区分分けした。KM 曲線を図 13 に示す。

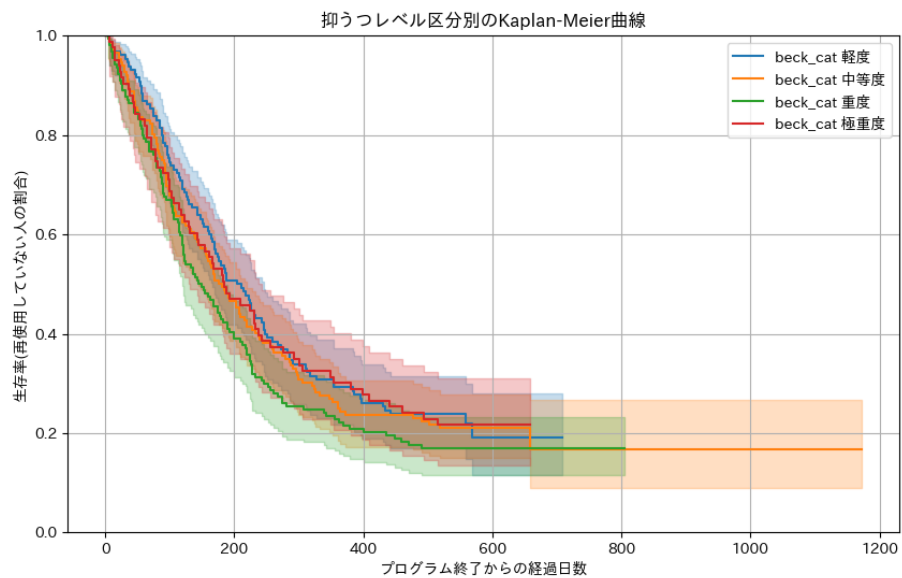


図 13. 抑うつレベル区分の Kaplan-Meier 曲線

上記の図を見ると、KM 曲線が多くの箇所で相互に交差していることがわかる。よって、重度, 極重度とそれ以外の 2 分で KM 曲線を図 14 に図示することとする。

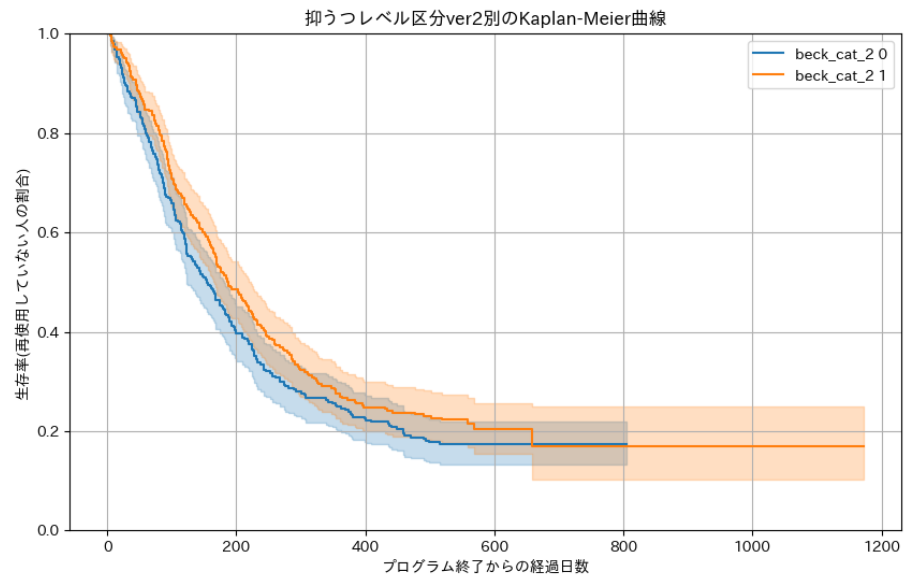


図 14. 抑うつレベル区分 ver2 の Kaplan-Meier 曲線

上記の図を見ると、曲線同士が交差していない。よって、抑うつレベル区分が再使用率に対して何らかの影響を与えていると推測する。そのため、ログランク検定を用いて有意性を判断する。

表 13. ログランク検定結果

	test_statistic	p	-log2(p)
0	3.248482	0.07149	3.806122

上記を見ると、p 値が有意水準 5%を上回っている。したがって、抑うつレベルが有意な影響を与えているとは言い切れないと評価する。

(ウ)Cox 回帰モデル

Kaplan-Meier 曲線とログランク検定結果をもとに Cox 回帰モデルにフィッティングする。Cox 回帰モデルは、比例ハザードモデルとも言い比例ハザード性の確認が必要である。そのため、Cox 回帰モデルのフィッティングを行ったうえで、Schoenfeld 残差を確認し比例ハザード性の仮定を確認することとする。Cox 回帰モデルのモデル式は、説明変数ベクトルを $X = (X_1, \dots, X_p)$ とすると以下ようになる。

$$h(t | X) = h_0(t)\exp(\beta_1 X_1 + \dots + \beta_p X_p)$$

まず、各列を一つずつ説明変数として単変量 Cox 回帰モデルを実施する。以下に単変量 Cox 回帰モデル結果を示す。

```
=== Univariate Cox model for 'heroic_cat' ===
```

	coef	exp(coef)	se(coef)	coef lower 95%	coef upper 95%	
covariate						
heroic_cat	0.252438	1.287159	0.095325	0.065603	0.439272	

	exp(coef)	lower 95%	upper 95%	cmp to	z	
covariate						
heroic_cat	1.067803	1.551578	0.0	2.648168		

	p	-log2(p)	
covariate			
heroic_cat	0.008093	6.949121	

図 15. 単変量(ヘロイン,コカイン)Cox 回帰モデル結果

上記の図は、ヘロイン、コカイン使用に関するカテゴリ変数を説明変数としてCox 回帰モデルである。これを見ると、Wald 検定において p 値が 0.008 となり有意水準 5%以下である。また、 $\exp(\text{coef})$ が 1.28 でありヘロインを使用している方が、していない人に比べて 1.28 倍再使用率が高いと評価できる。ただし、比例ハザード性について検定し 1.28 倍という比率が時間に依存しないか確認する必要がある。よって、Schoenfeld 残差を実施し結果を図 16,17 に示す。

```
--- Schoenfeld residuals PH test ---
```

time_transform	rank		
null_distribution	chi squared		
degrees_of_freedom	1		
model	<lifelines.CoxPHFitter: fitted with 575 total ...		
test_name	proportional_hazard_test		
test_statistic	p	-log2(p)	
heroic_cat	0.03	0.87	0.20

図 16. Schoenfeld 残渣結果

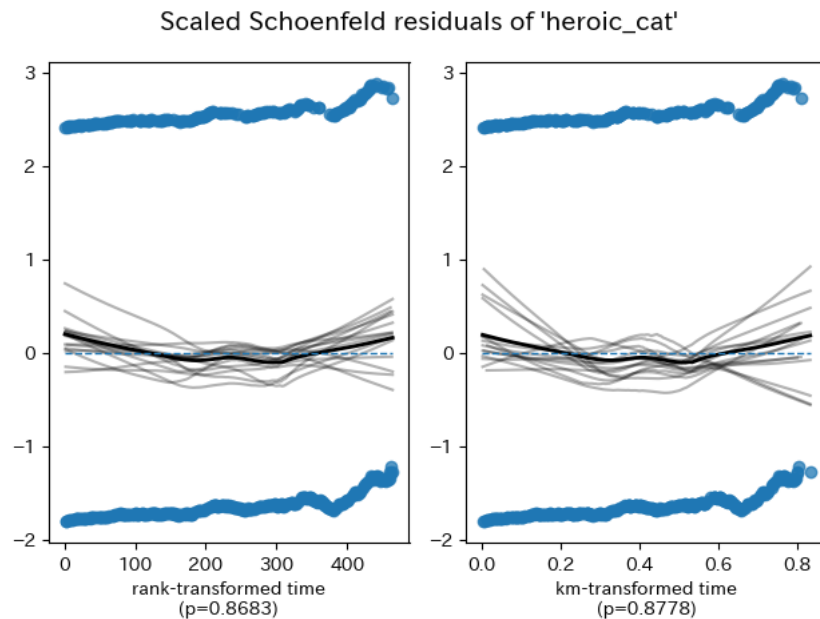


図 17. Schoenfeld 残渣プロット

図 16,17 を見ると、p 値が 0.05 を上回っており帰無仮説を棄却できない。つまり、比例ハザード性がなく時間に依存すると言いきれない。そのため、Cox 回帰モデルとしての比例ハザード仮定が成り立っていると評価できる。

次に、薬物注射の直近利用に関する Cox 回帰モデルについて注目する。z ウ 18 は、そのモデルの結果である。これを見ると、p 値が 0.0005 であり有意水準 5%を下回っており $\exp(\text{coef})$ が 0.72 となるため直近に使用歴がある ($\text{ivhx_cat}=0$) のほうが、1.4 倍(0.72 の逆数)再使用率が高くなるといえる。これも同様に、Schoenfeld 残渣を実施し比例ハザード性の仮定を確認する。図 19, 20 を見ると、p 値が 0.82 で残渣プロットも大きな変化が見られないため、比例ハザード性があると統計的に評価できる。

```

=== Univariate Cox model for 'ivhx_cat' ===
      |      |      |      |      |      |      |      |      |
covariate.....coef exp(coef) se(coef) coef lower 95% coef upper 95% \
ivhx_cat -0.323829  0.723374  0.093457      -0.507      -0.140657...

      |      |      |      |      |      |      |      |      |
covariate.....exp(coef) lower 95% exp(coef) upper 95% cmp to      z      p \
ivhx_cat      0.602299      0.868787      0.0      -3.46501  0.00053...

      |      |      |      |      |      |      |      |      |
covariate.....-log2(p).....
ivhx_cat    10.881145 ..

```

図 18. 単変量(薬物注射)Cox 回帰モデル結果

--- Schoenfeld residuals PH test ---			
time_transform	rank		
null_distribution	chi squared		
degrees_of_freedom	1		
model	<lifelines.CoxPHFitter: fitted with 575 total ...		
test_name	proportional_hazard_test		
test_statistic	p	-log2(p)	
ivhx_cat	0.05	0.82	0.28

図 19. Schoenfeld 残渣結果

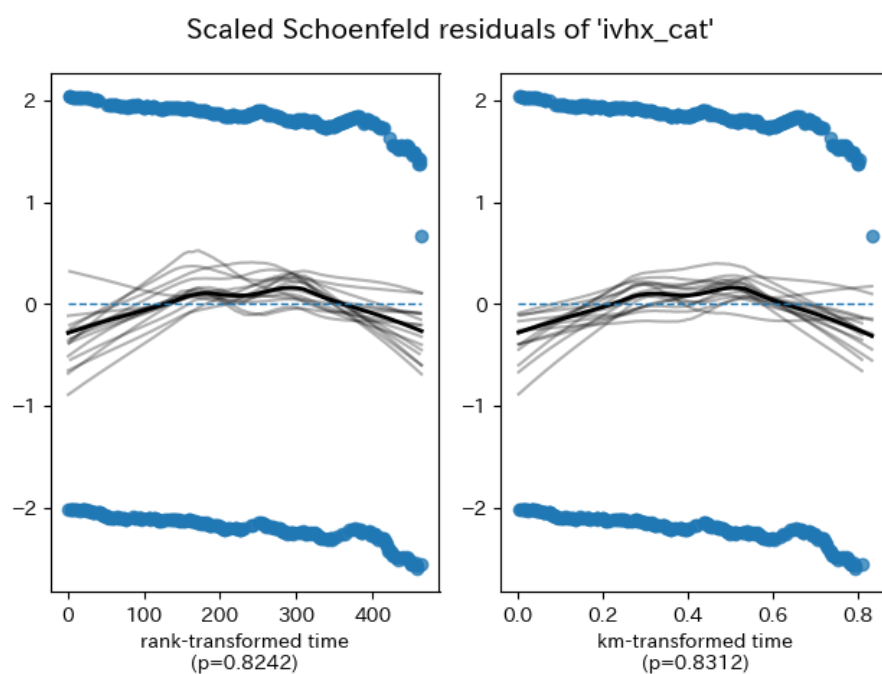


図 20. Schoenfeld 残渣結果

次に、人種に注目する。人種では白人が 0, その他が 1 であった。この時、Cox 回帰モデルの説明変数として利用した結果を図 21 に示す。図 21 では、p 値が 0.007 になり $\exp(\text{coef})$ が 0.74 となっている。よって、白人のほうが 1.35 倍再使用率が高くなると統計的に評価できる。そのため、Schoenfeld 残渣を用いて比例ハザード仮定を崩していないか図 22,23 に示して確認する。

```

=== Univariate Cox model for 'fac_race' ===

```

	coef	exp(coef)	se(coef)	coef lower 95%	coef upper 95%	\
covariate						
fac_race	-0.299414	0.741253	0.110323	-0.515643	-0.083184	...

	exp(coef)	lower 95%	upper 95%	cmp to	z	\
covariate						
fac_race	0.597116		0.920182	0.0	-2.71397	...

	p	-log2(p)
covariate		
fac_race	0.006648	7.232815

図 21. 単変量(人種)Cox 回帰モデル結果

```

--- Schoenfeld residuals PH test ---

```

time_transform	rank		
null_distribution	chi squared		
degrees_of_freedom	1		
model	<lifelines.CoxPHFitter: fitted with 575 total ...		
test_name	proportional_hazard_test		
test_statistic	p	-log2(p)	
fac_race	1.33	0.25	2.01

図 22. Schoenfeld 残渣結果

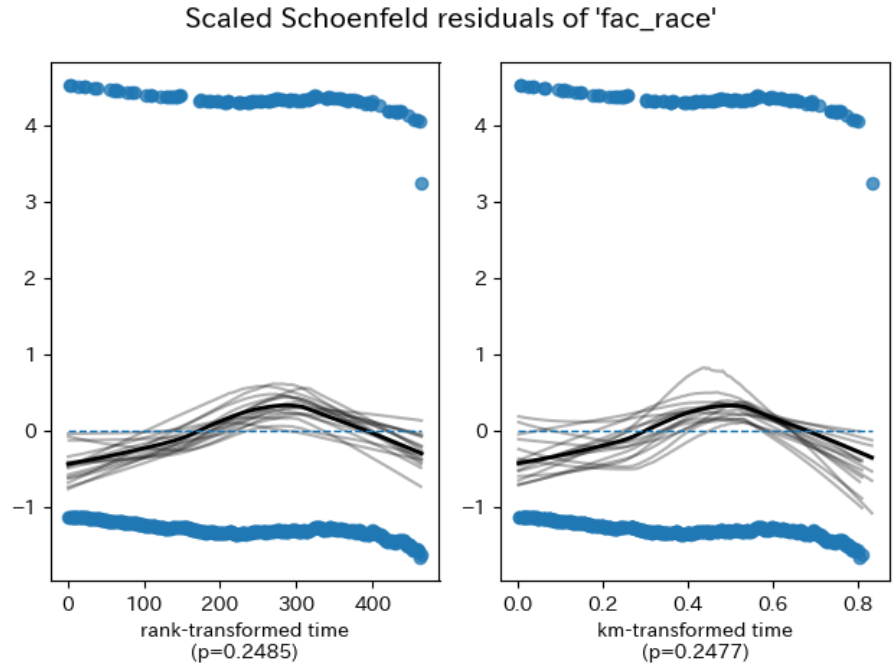


図 23. Schoenfeld 残渣プロット

図 22,23 を見ると、p 値が 0.25 で残渣プロットが明らかに曲がっているとは言えない点より、比例ハザード性があると統計的に評価できる。

次に、治療期間について、単変量 Cox 回帰を実施する。結果を図 24 に示す。図 24 を見ると、p 値が 0.01 となっており $\exp(\text{coef})$ が 0.78 となっている。したがって、治療期間が長いと 0.78 倍再使用率が增加する。要するに、治療期間を長くすることで再使用率を下げるができることといえる。本回帰モデルについて、仮定として比例ハザード性がある場合のみ評価できるため、Schoenfeld 残渣を図 25,26 に示す。それらの図を見ると、p 値が 0.07 であるためわずかに有意水準を超えている。この点について、治療プログラムの長さは、短期的には影響するが長期的には再使用率に大きな影響を与えないのではないかと考える。

```
=== Univariate Cox model for 'fac_treat' ===
      |      |      |      |      |      |      |      |      |      |
      |      |      |      |      |      |      |      |      |      |
covariate.....coef exp(coef) se(coef) coef lower 95% coef upper 95% \
fac_treat -0.237443  0.788642  0.093084      -0.419884      -0.055001...

      |      |      |      |      |      |      |      |      |      |
      |      |      |      |      |      |      |      |      |      |
covariate.....exp(coef) lower 95% exp(coef) upper 95% cmp to      z \
fac_treat      0.657123      0.946484      0.0 -2.550845...

      |      |      |      |      |      |      |      |      |      |
      |      |      |      |      |      |      |      |      |      |
covariate.....p -log2(p).....
fac_treat  0.010746  6.540027..
```

図 24 単変量(治療期間)Cox 回帰モデル結果

```
--- Schoenfeld residuals PH test ---

      time_transform      rank
null_distribution      chi squared
degrees_of_freedom      1
model <lifelines.CoxPHFitter: fitted with 575 total ...
test_name      proportional_hazard_test
test_statistic      p      -log2(p)
fac_treat      3.26  0.07      3.82
```

図 25. Schoenfeld 残渣結果

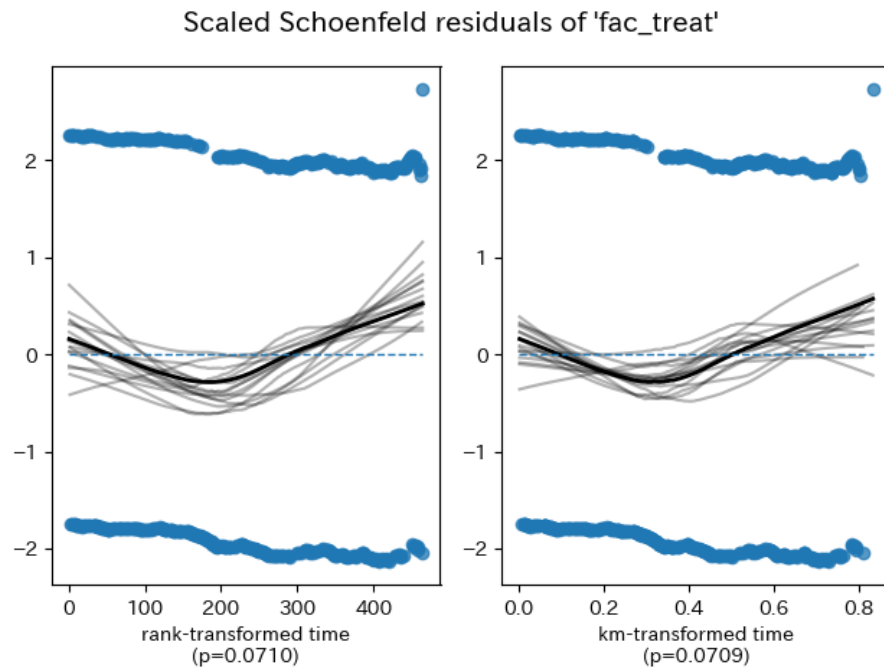


図 26. 残渣プロット

次に、治療施設について注目する。Cox 回帰モデル結果の図 27 を見ると、p 値が 0.29 となり有意水準 5%を大きく上回っている。そのため、治療施設の差が生存関数に有意な影響を与えているとは言い切れないと評価する。よって、多変量 Cox 回帰モデルでは、説明変数から除外することとする。

```

=== Univariate Cox model for 'fac_site' ===

```

	coef	exp(coef)	se(coef)	coef lower 95%	coef upper 95%	\
covariate:						
fac_site	-0.10647	0.899002	0.101678	-0.305756	0.092816	...

	exp(coef)	lower 95%	upper 95%	cmp to	z	\
covariate:						
fac_site	0.736567		1.09726	0.0	-1.047125	...

	p	-log2(p)
covariate:		
fac_site	0.295042	1.761008

図 27. 単変量(治療施設)Cox 回帰モデル結果

次に、年齢区分に関する Cox 回帰モデルに注目する。結果を図 28 に示す。以下を見ると、p 値が僅かに有意水準 5%を下回っている。よって、本解析基準に従うと有意な影響を与えていると評価できる。また、exp(coef)について、1.20 となっているため、年齢が中央値(30 歳前半)より高い方が、1.2 倍

再使用率が高くなると評価できる。よって、Scoenfeld 残渣を用いて仮定を満たしているか確認する。残渣結果を図 29.30 に示す。

```
=== Univariate Cox model for 'age_cat' ===
```

	coef	exp(coef)	se(coef)	coef lower 95%	\
covariate					
C(age_cat)[T.young]	0.183569	1.201498	0.093106	0.001084	

	coef upper 95%	exp(coef) lower 95%	exp(coef) upper 95%
covariate			
C(age_cat)[T.young]	0.366055	1.001084	1.442034

	cmp to	z	p	-log2(p)
covariate				
C(age_cat)[T.young]	0.0	1.971606	0.048655	4.361278

図 28. 単変量(年齢カテゴリ)Cox 回帰モデル結果

```

--- Schoenfeld residuals PH test ---

      time_transform                                rank
null_distribution                                chi squared
degrees_of_freedom                                1
      model  <lifelines.CoxPHFitter: fitted with 575 total ...
test_name                                proportional_hazard_test

      test_statistic      p  -log2(p)
C(age_cat)[T.young]      0.91  0.34    1.56

```

図 29. Schoenfeld 残渣結果

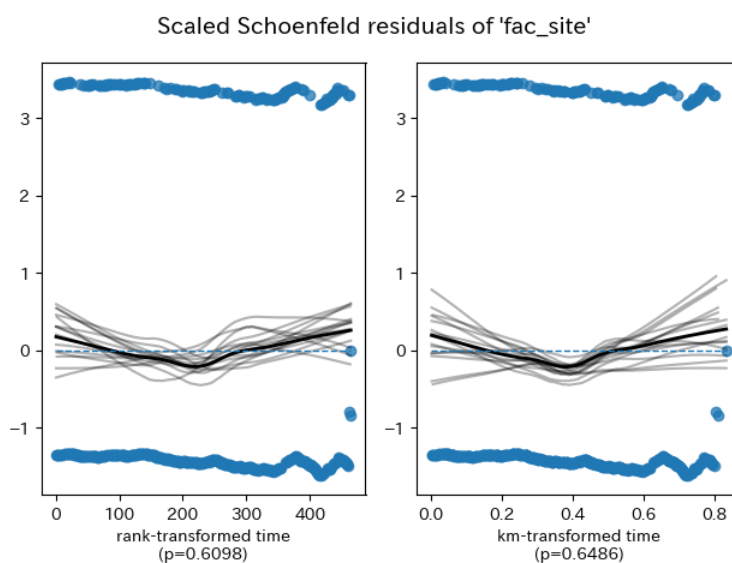


図 30. 残渣プロット

上記を見ると、p 値が 0.34 であり、比例ハザード性を満たしていると評価できる。よって、本モデルによって得られた年齢は有意な影響を与えているといえる。

これらより、比例ハザード性を持つ変数を取り入れた多変量 Cox 回帰モデルを作成する。図 31 に回帰結果を示す。

```
=== Coefficients & Wald test ===
               coef  exp(coef)  se(coef)      z      p
covariate
heroic_cat      0.085290    1.089032  0.128683  0.662789  0.507466
ivhx_cat       -0.290115    0.748177  0.127839 -2.269373  0.023246
fac_race       -0.250450    0.778450  0.111710 -2.241972  0.024963
C(age_cat)[T.young] 0.264451    1.302716  0.095500  2.769125  0.005621

=== Likelihood ratio test ===
 $\chi^2 = 24.956$ , df = 4, p = 0.0001
```

図 31. 多変量 Cox 回帰結果

上記の結果を見ると、LRT の尤度比推定では有意水準を下回る p 値を得られたが、各説明変数の p 値についてヘロイン,コカインの使用履歴について 5%を上回ってしまった。そのため、ヘロイン,コカイン使用歴についての説明変数を除外して再度モデリングをすることとする。モデリング結果を図 32 に示す。

```
=== Coefficients & Wald test ===
               coef  exp(coef)  se(coef)      z      p
covariate
ivhx_cat       -0.345081    0.708163  0.096718 -3.567891  0.000360
fac_race       -0.248341    0.780094  0.111653 -2.224216  0.026134
C(age_cat)[T.young] 0.260333    1.297362  0.095264  2.732748  0.006281

=== Likelihood ratio test ===
 $\chi^2 = 24.516$ , df = 3, p = 0.0000
```

図 32. 多変量 Cox 回帰結果 2

上記の図を見ると、各説明変数において Wald 検定の p 値も、LRT 尤度比検定も、どちらも有意水準 5%を下回っている。よって、比例ハザード性を満たしていることも踏まえると本モデルに使用している説明変数は、どれも有意な影響を薬物再使用に対して与えていると考える。

そして、具体的なモデルの結果としては多変量 Cox 回帰モデルの尤度比検定

において、 $\chi^2=24.516$ 、 $df=3$ 、 $p<0.0001$ であった。これは ivhx_cat、fac_race、age_cat の 3 変数が無説明変数モデルより有意に当てはまりを改善したことを示している。ivhx_cat（コカイン使用歴）の係数は-0.345、Wald 検定 $p=0.00036$ であった。exp(coef)=0.708 というハザード比は、コカイン未使用群のリスクが使用群の約 0.71 倍であることを意味する。すなわち、コカイン使用歴はイベント発生リスクの上昇と関連していると評価できる。次に、fac_race（人種）の係数は-0.248、 $p=0.0261$ であった。ハザード比 0.780 は白人でないグループでリスクが約 0.78 倍であることを示している。この結果は人種が再使用リスクに独立した影響を及ぼす可能性を示唆する。最後に、age_cat[T.young]（若年群）は係数 0.260、 $p=0.00628$ であった。exp(coef)=1.297 のハザード比は若年群のリスクが約 1.30 倍高いことを示している。年齢が若いほどイベントリスクが増加するという逆直感的な結果となった。

各係数の標準誤差（約 0.09~0.11）は比較的小さく、推定の精度が高い。全体のカイ二乗統計量はモデルの予測力が強いことを示している。Wald 検定によって各変数の効果が偶然ではないことが確認された。説明変数の影響方向は変数ごとに異なり、保護効果とリスク効果の両方が見られると評価できる。

一方で、比例ハザード性の前提を満たしているかの検証も重要である。そこで、複数の因子の交互作用項の導入によってさらに複雑な関係を調べる必要がある。さらに、外部コホートによる検証でモデルの頑健性を評価すべきである。臨床的にはヘロイン使用歴や年齢、人種情報を考慮したリスク層別が有用となる可能性があると結論付ける。最終的に本モデルは薬物再使用者予測における主要リスク因子を示す有益な指標となると評価する。

4. 解析プログラム

今回、使用したコードはすべて GitHub に公開している。

(https://github.com/yut0takagi/Ass_ReliabilityEngineering)

5. 参考文献

1. YahooNews, “ヤバすぎる合成麻薬「フェンタニル」日本とばっちり 中国から「経由地」にされ米国へ？ トランプ関税に影響？”,
(<https://news.yahoo.co.jp/articles/e133b8320e8383772bdb1f870cff88619ba03c70>), 2025/07/15 参照。
2. David Nutt, Leslie A King, William Saulsbury, Colin Blakemore,
“Development of a rational scale to assess the harm of drugs of potential misuse”
2025/07/16 参照。