

AMED生物統計家育成推進事業主催シンポジウム 2021/12/24@オンライン
「行動変容プログラム/アプリの開発と評価」
～産官学における取り組みと事例に基づく議論～

行動変容アプリの臨床試験における 統計的アプローチ

東京大学大学院情報学環・生物統計情報学講座

特任講師 上村鋼平

本日の内容

- はじめに
- 海外の開発・審査体制
- 統計的なアプローチ

はじめに

■ 治療用アプリ

- ◆医薬品医療機器等法(薬機法) にもとづく薬事承認を受けた医療機器に該当し、医師が処方して、患者が日常生活で用いるアプリ
- ◆2014年の薬機法改正を機に、ソフトウェアが「プログラム医療機器」として扱うようになったことが背景

■ ヘルスケア関連のアプリ

- ◆一般のアプリと同様、自由に利用可能

治療用と非治療用の違いは？

■ アプリの位置づけの違い

- ◆ 単なる健康管理であれば、非治療用
- ◆ **治療**もしくは**治療の補助**であれば、治療用

■ 治療用のアプリとして薬事承認・保険適用を目指すなら

- ◆ **適切な臨床試験デザイン**に基づく、治療用アプリとしての**有効性**と**安全性**に関する**エビデンスの構築**が**必須**

※ 非治療用も臨床試験やっています。

治療用アプリはなぜ必要か

■ 服薬による治療の限界

- ◆生活習慣病などは明確な症状が現れず、**患者の病識が乏しい**

 - ✓ニコチン依存症では禁煙を継続できるのは3割以下

 - ✓糖尿病では継続治療率は5割強といわれている

- ◆従来の治療法の**限界を補う位置づけ**として治療用アプリが登場

■ 例えば、治療アプリにより患者が体調などの情報を入力

- ◆アプリがエビデンスに基づく最新情報により、**適切なタイミングで個別に最適化された診療的アドバイス**を患者に提供

- ◆次回の受診までの治療空白期間を埋め、治療継続意欲を持続

きめ細やかなケアにより生存が延びた例

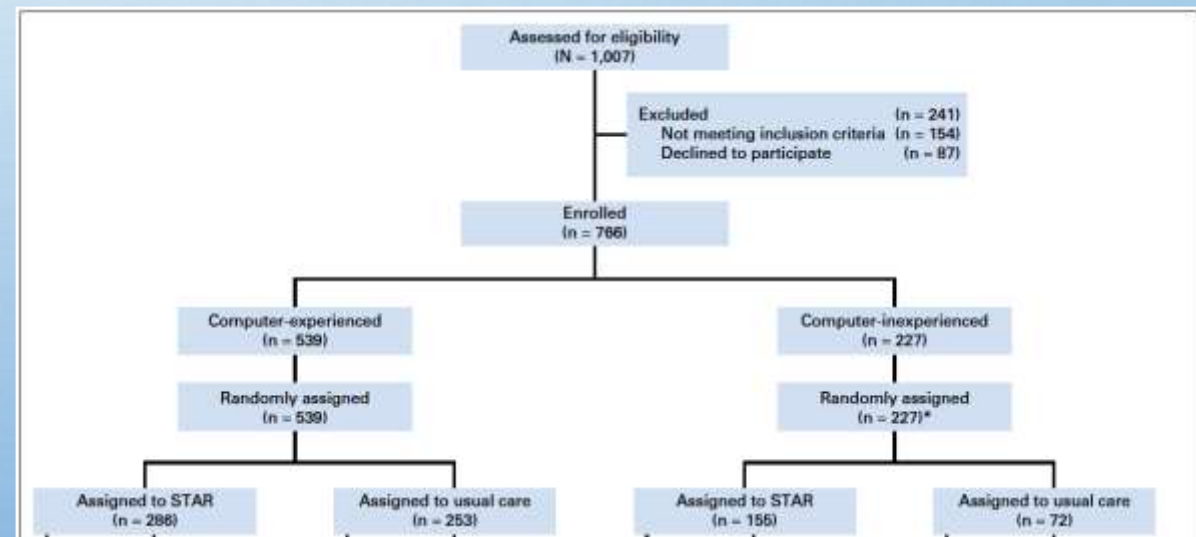
■ 進行固形がん患者766例対象

◆ Usual Care群: 医師の問診

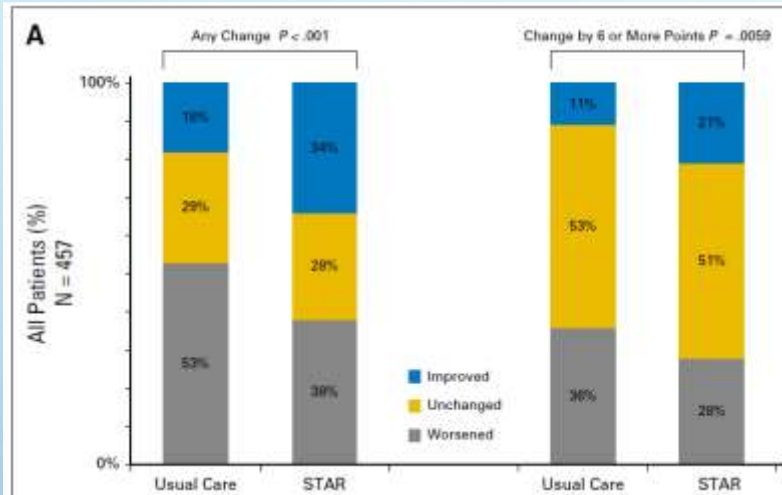
◆ **STAR**群: タブレット端末を使用し、患者の症状をモニタリング

✓ 症状が悪化すると自動e-mailが看護師へ送信されるシステム

■ QOLや生存を比較するRCT



Results



- 主要評価項目: 6カ月後のQOLの変化量
 - ◆ 症状モニタリング(STAR)群の方がQOLが改善
⇒ **34% vs. 18% (改善割合), $p < .001$**

- 副次評価項目: 生存期間
 - ◆ 症状モニタリング(STAR)群の方が生存期間が延長傾向
 - ◆ ⇒ **75.1% vs. 68.6% (1年生存率), $p = .05$**

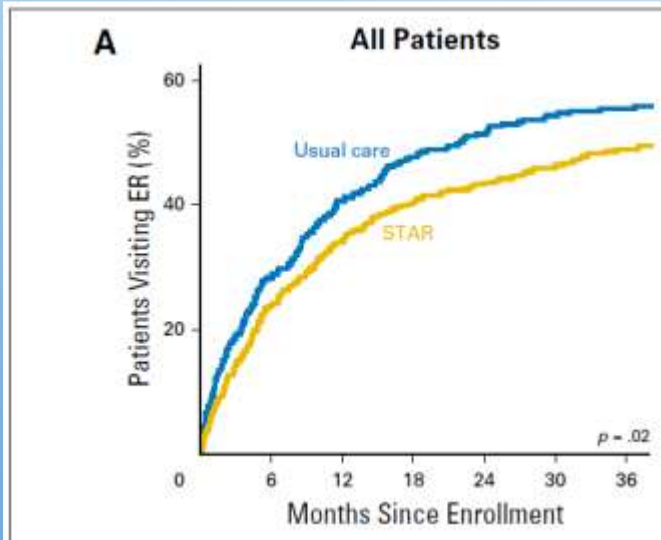


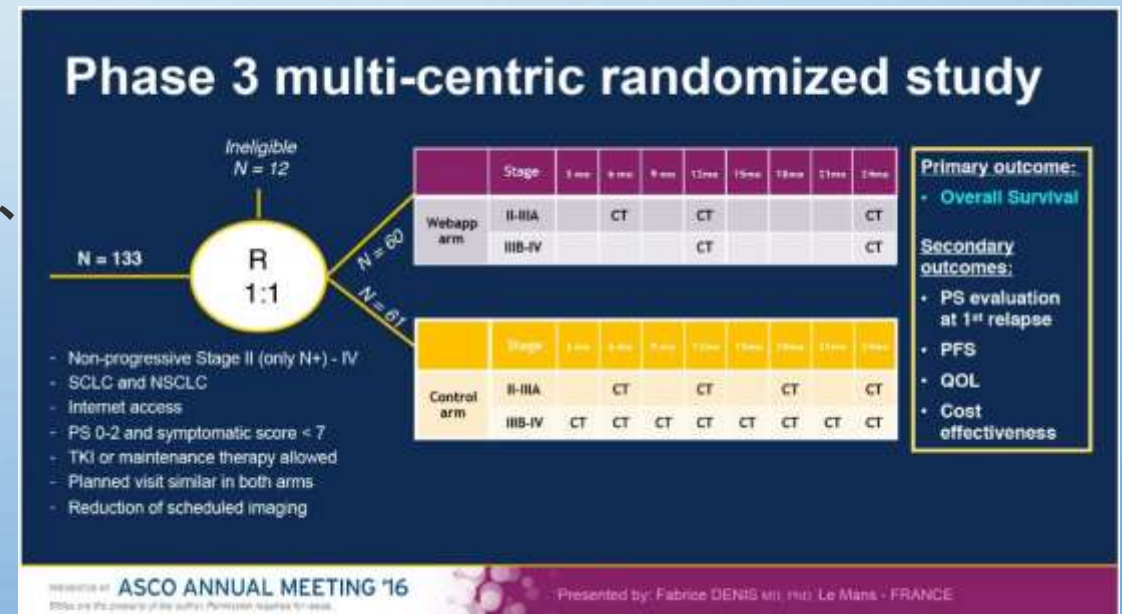
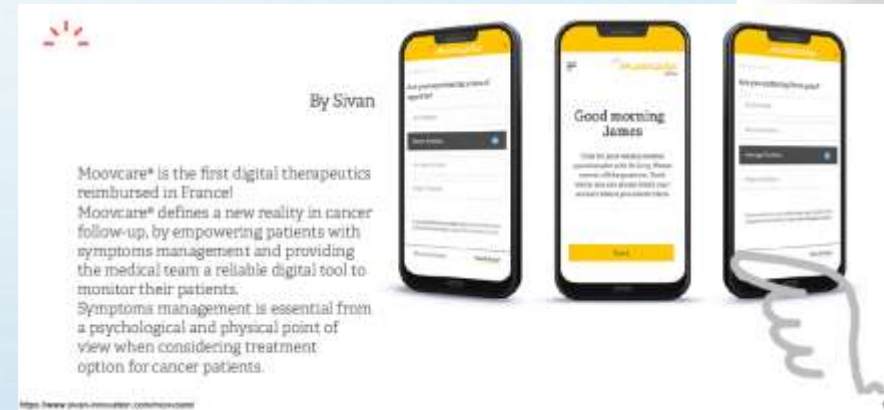
Table 3. Overall and Quality-Adjusted Survival at 12 Months

Patients	N	STAR (95% CI)	Usual Care (95% CI)	P (Univariable)*	P (Multivariable)*
Overall survival, % alive at 1 year					
All patients	766	75.1 (70.7 to 79.0)	68.6 (63.2 to 73.6)	.03	.05
Subgroup analysis, % alive at 1 year					
Computer inexperienced	227	74.2 (66.6 to 80.9)	59.7 (47.5 to 71.1)	.03	.02
Computer experienced	539	75.5 (70.1 to 80.4)	71.1 (65.1 to 76.7)	.25	.45
Quality-adjusted 12-month survival, months					
All patients	757†	8.7 (8.3 to 9.0)	8.0 (7.6 to 8.4)	.002	.004
Subgroup analysis, months					
Computer inexperienced	220†	8.3 (7.8 to 8.8)	7.2 (6.3 to 8.2)	.03	.02
Computer experienced	537†	8.8 (8.5 to 9.2)	8.2 (7.7 to 8.6)	.02	.046

Abbreviation: STAR, Symptom Tracking and Reporting web-based self-reporting system (study intervention).
 * P values for between-arm comparisons. Multivariable analyses controlled for age, sex, cancer type, race, and education level. For overall analyses, subgroup assignment (computer experienced or computer inexperienced) was also included as a covariate.
 †Participants with missing baseline health-related quality of life scores not included in quality-adjusted survival analysis.

Moovcare® ASCO2016で注目演題

- 初回の化学療法、放射線療法、外科手術を終えた**ステージ3～4期の肺がん患者133人**をアプリを介した経過観察群と、標準的な経過観察群に割り付け
- 患者はアプリを用いて、毎週12の症状について自己評価を行う
- **アプリは12の症状を解析し担当医に報告**
- アルゴリズムに則して症状の変化を評価し、医師に注意喚起のEメールを送信
- **医師は検査や診察の必要性を確認し、患者にフィードバック**

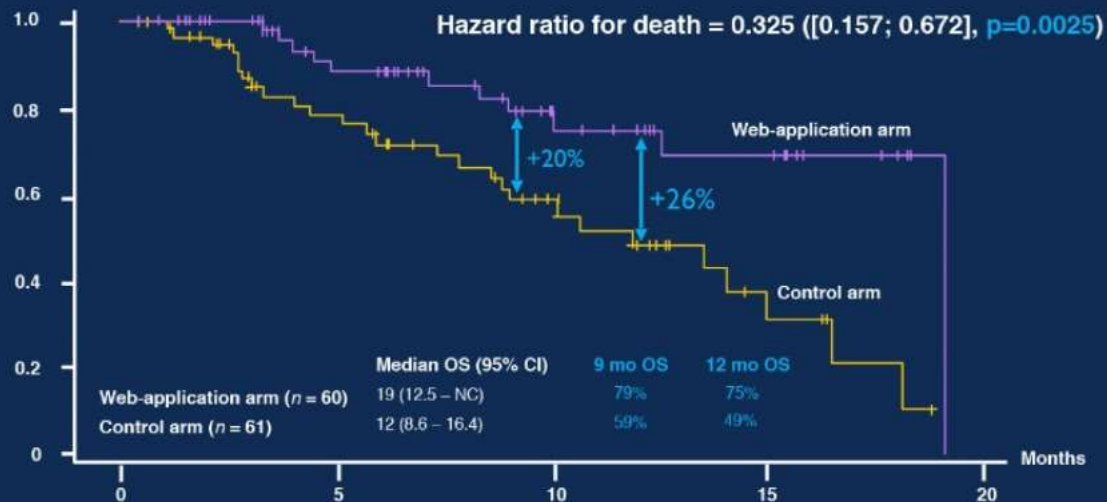


中間解析により有効中止: 想定以上の差を意味

- OS中央値: アプリ群**19カ月** vs. Control群**12カ月**
◆ **HR[95%CI]: 0.325[0.157, 0.672], p=0.0025**
- 1年生存率: アプリ群**75%** vs. Control群**49%**

アプリが治療自体を担ったわけではないが、自分が患者ならこのシステムは導入して欲しい

Overall Survival Improvement



PRESENTED AT: ASCO ANNUAL MEETING '16
Slides are the property of the author. Permission required for reuse.

Presented by: Fabrice DENIS MD, PhD, Le Mans - FRANCE

■ QOLも改善傾向

Quality of Life (mean score at 6 months)



PRESENTED AT: ASCO ANNUAL MEETING '16
Slides are the property of the author. Permission required for reuse.

Presented by: Fabrice DENIS MD, PhD, Le Mans - FRANCE

海外の 開発・審査体制

どんなエビデンスに
基づき審査され、
どんな制度が運用
されているか？



治療用アプリの世界初の承認事例@FDA

2型糖尿病患者指導アプリ: BlueStar®

Cluster-Randomized Trial of a Mobile Phone Personalized Behavioral Intervention for Blood Glucose Control

CHARLENE C. QUINN, RN, PhD
MICHELLE D. SHARRETT, PhD
MICHAEL L. TERRIN, MD, MPH

ERIC A. BARR, BA
SUDHANA H. BALLEW, BA
ANN L. GRUBER-BALDING, PhD

interventions to assist patients and PCPs (12–14). The Mobile Diabetes Intervention Study, reported here, evaluated a diabetes coaching system, using mobile phones as patient/provider portals for patient-specific treatment and communication. The hypothesis tested was that mobile telephone feedback on self-management of blood glucose

OBJECTIVE—To test whether adding mobile application coaching and patient/provider web portals to community primary care compared with standard diabetes management would reduce glycosylated hemoglobin levels in patients with type 2 diabetes.

■ どのような臨床試験デザインが採用され、どんな結果だったか

- アカデミア所属のない医師が3名以上いる、プライマリケア施設を対象として施設単位のCluster RCT
- 通常ケアと介入3グループへ1.5:1:1:1.5で割付
26施設が9, 4, 6, 7施設へランダム化
 - ◆ 通常ケア (Group 1): 56例
 - ◆ 介入群 (Group 2~4): 23, 22, 62例

アプリを利用した介入内容

- Group 1: 通常どおりのケア (Control)
- Group 2: 血糖測定器と同期されたアプリが患者へ自動送信する指導・動機づけメッセージ
- Group 3: Group 2 + 医師に患者の生データのみ送信
- Group 4: Group 2 + 医師に患者データの分析結果 (血糖・メタボリック管理状況、服薬、自己管理状態、治療ガイドラインのエビデンスに紐づいた判断サポート) を送信
 - ◆ 医師には、インセンティブの支払い (250\$/1患者)
 - ◆ 医師は、通常ケア群か介入群かどうかは非盲検

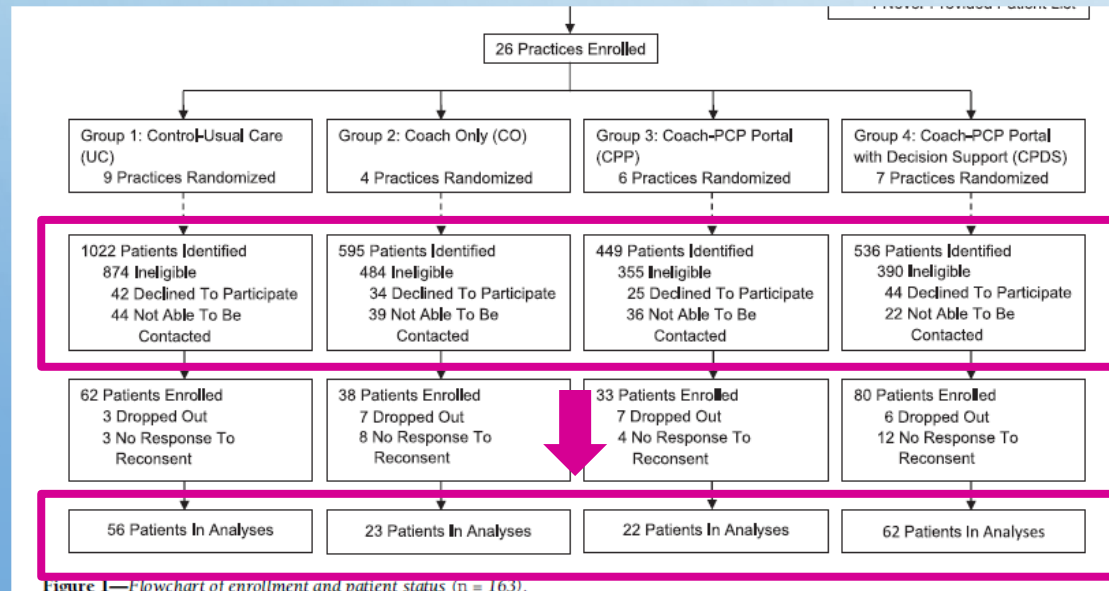


Figure 1—Flowchart of enrollment and patient status (n = 163).

患者背景に偏りはある程度ある

- HbA1cが9%以上の重症者が4割弱～5割以上
- ◆ BMIの平均値は35kg/m², 40以上が2～3割以上
- ◆ 高血圧は5～8割、高脂血症は5～6割

Table 1—Baseline characteristics of patients and primary and secondary study outcomes

	Group 1: UC (n = 56)		Group 2: CO (n = 23)		Group 3: CPP (n = 22)		Group 4: CPDS (n = 62)	
	n	% or mean ± SD*	n	% or mean ± SD*	n	% or mean ± SD*	n	% or mean ± SD*
Baseline characteristics								
Glycated hemoglobin (%)	56	9.2 ± 1.7	23	9.3 ± 1.8	22	9.0 ± 1.8	62	9.9 ± 2.1
7.5–8.9	35	62.5	13	56.5	13	59.1	28	45.2
≥9	21	37.5	10	43.5	9	40.9	34	54.8
Age (years)	56	53.2 ± 8.4	23	52.8 ± 8.0	22	53.7 ± 8.2	62	52 ± 8.0
Sex								
Male	28	50	12	52.2	10	45.5	31	50
Female	28	50	11	47.8	12	54.5	31	50
BMI								
BMI (kg/m ²)	56	34.3 ± 6.3	23	36.9 ± 7.5	22	35.5 ± 10.3	62	35.8 ± 7.1
Underweight (16.5–18.4 kg/m ²)	1	1.8	0	0	1	4.5	0	0
Normal (18.5–24.9 kg/m ²)	0	0	0	0	2	9.1	1	1.6
Pre-obese (25–29.9 kg/m ²)	11	19.6	4	17.4	7	31.8	12	19.4
Obese class 1 (30–34.9 kg/m ²)	22	39.3	6	26.1	1	4.5	18	29
Obese class 2 (35–39.9 kg/m ²)	10	17.9	5	21.7	3	13.6	17	27.4
Obese class 3 (≥40 kg/m ²)	12	21.4	8	34.8	8	36.4	14	22.6
Comorbidities								
Hypertension	29	51.8	18	78.3	13	59.1	43	69.4
Hypercholesterolemia	34	60.7	11	47.8	14	63.6	36	58.1
Coronary artery disease	5	8.9	2	8.7	0	0	5	8.1
Microvascular complications, any	8	14.3	1	4.3	2	9.1	6	9.7

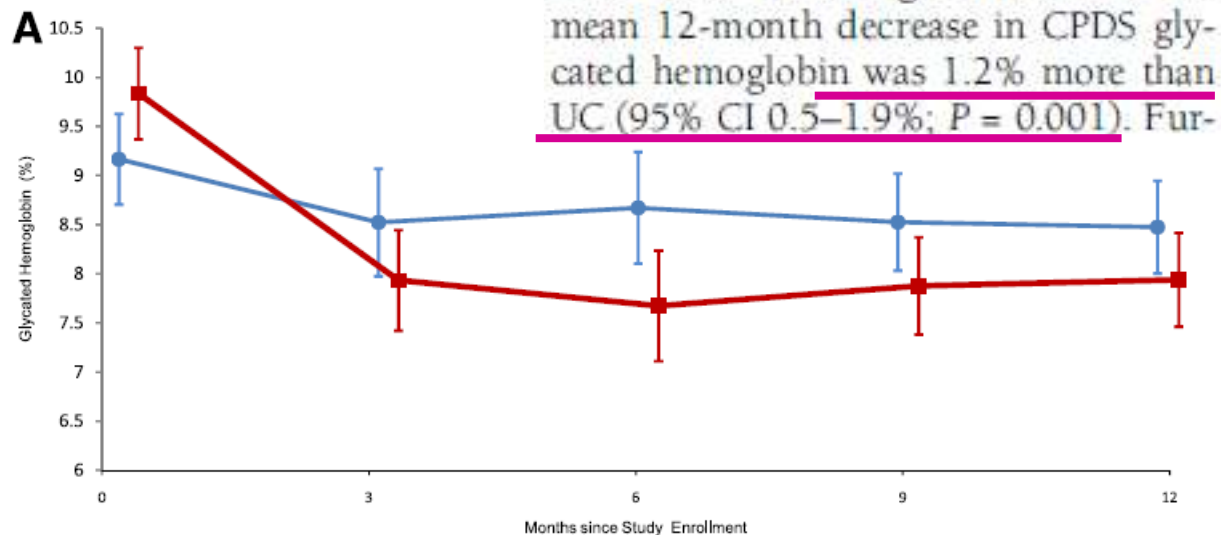
■ 主要評価項目の結果

Both had greater 12-month glycated hemoglobin reductions than UC (CO, $P = 0.027$; CPP, $P = 0.40$). CO and CPP were similar to CPDS over all follow-up time points ($P > 0.05$ for both comparisons).

Table 1—Continued

	Group 1: UC (n = 56)		Group 2: CO (n = 23)		Group 3: CPP (n = 22)		Group 4: CPDS (n = 62)	
	n	% or mean $\pm$ SD*	n	% or mean $\pm$ SD*	n	% or mean $\pm$ SD*	n	% or mean $\pm$ SD*
Primary outcome: glycated hemoglobin (%)†								
Baseline	56	9.2 $\pm$ 1.7	23	9.3 $\pm$ 1.8	22	9.0 $\pm$ 1.8	62	9.9 $\pm$ 2.1
3 Months	30	8.2 $\pm$ 1.2	13	7.6 $\pm$ 1.2	9	7.5 $\pm$ 0.6	41	7.8 $\pm$ 1.3
6 Months	27	8.6 $\pm$ 2.0	15	7.6 $\pm$ 1.1	11	7.6 $\pm$ 0.7	30	7.5 $\pm$ 1.2
9 Months	23	8.2 $\pm$ 1.3	13	7.7 $\pm$ 0.9	7	7.6 $\pm$ 0.8	32	7.7 $\pm$ 2.1
12 Months	51	8.5 $\pm$ 1.8	21	7.7 $\pm$ 1.0	21	7.9 $\pm$ 1.4	56	7.9 $\pm$ 1.7
Change from baseline to 12 months (mean)‡		-0.7		-1.6		-1.2		-1.9
Change from baseline to 12 months (95% CI)‡		-1.1 to -0.3		-2.3 to -1.0		-1.8 to -0.5		-2.3 to -1.5

Randomized trial of mobile diabetes coaching



■ サブグループ解析の結果

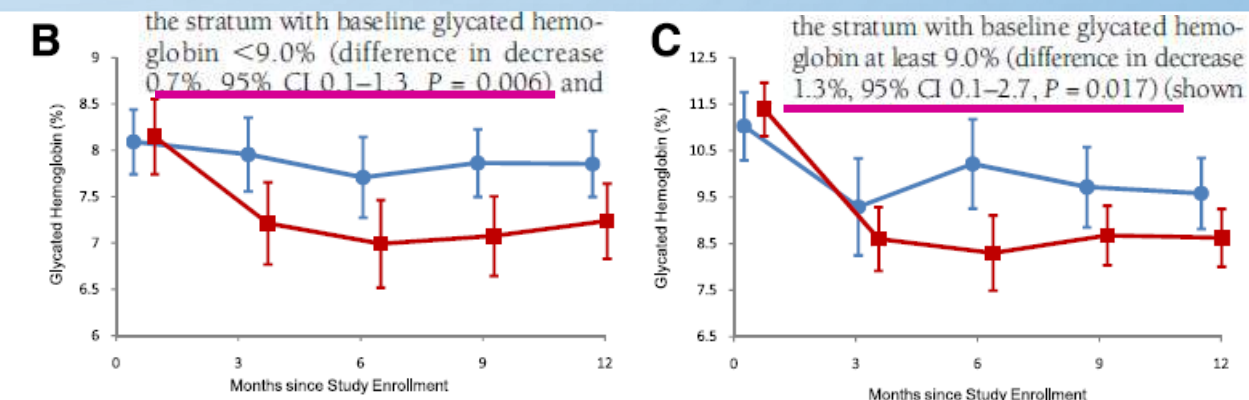


Figure 2—Primary study outcome and baseline A1C stratified analyses.

副次評価項目の結果

■ 収縮期血圧、LDLコレステロールは

◆ベースラインと比較しても、統計的に有意な低下は認められず

◆群間でも統計的に有意な差はない

Secondary outcomes, laboratory values

Systolic blood pressure (mmHg)

Baseline	56	130 ± 22	23	130 ± 18	22	133 ± 14	62	130 ± 14
12 Months	45	133 ± 20	21	134 ± 25	20	134 ± 16	51	128 ± 19
Change from baseline to 12 months (mean)†		+2		+4		2		-2
Change from baseline to 12 months (95% CI)†		-3 to 7		-4 to 11		-6 to 10		-6 to 3

Diastolic blood pressure (mmHg)

Baseline	56	78 ± 12	23	79 ± 11	22	79 ± 9	62	79 ± 9
12 Months	45	79 ± 13	21	82 ± 11	20	78 ± 9	51	78 ± 10
Change from baseline to 12 months (mean)†		+1		+2		-2		-1
Change from baseline to 12 months (95% CI)†		-2 to 4		-2 to 7		-6 to 3		-4 to 2

LDL (mg/dL)

Baseline	51	102 ± 36	23	103 ± 29	22	103 ± 33	55	106 ± 33
12 Months	42	91 ± 34	19	94 ± 32	15	94 ± 47	45	102 ± 32
Change from baseline to 12 months (mean)†		-6		-8		-14		-5
Change baseline to 12 months (95% CI)†		-15 to 3		-21 to 5		-29 to 0		-13 to 4

※ 治療用アプリBlueStar[®]は、治療ではない

■ 約1か月で審査が終了

 DEPARTMENT OF HEALTH & HUMAN SERVICES
Public Health Service
Food and Drug Administration
10903 New Hampshire Avenue
Document Control Center - W
Silver Spring, MD 20993-0001

November 22, 2016

WellDoc, Incorporated
Danielle Dorfman
Manager, Regulatory Affairs and Quality System
10221 Wincopin Circle, Suite 150
Columbia, Maryland 21044

Re: K162225
Trade/Device Name: WellDoc Bluestar (WellDoc DiabetesManager[®] System and DiabetesManager[®]-Rx System)
Regulation Number: 21 CFR 880.5725
Regulation Name: Infusion Pump
Regulatory Class: II
Product Code: MRZ, LNX, NDC
Dated: October 26, 2016
Received: October 27, 2016

Feature	Subject Device
Indications for Use	DiabetesManager (OTC Use): The WellDoc DiabetesManager[®] System is indicated for use by healthcare providers (HCPs) and their adult patients - aged 21 years and older - who have type 2 diabetes. The DiabetesManager System is intended to provide secure capture, storage, and transmission of blood glucose data as well as information to aid in diabetes self-management. The DiabetesManager System analyzes and reports blood glucose test results and supports medication adherence. It includes software intended for use on mobile phones or personal computers in the home or in professional healthcare settings. The software also allows for entry of other diabetes-related healthcare information and provides educational information. The DiabetesManager System is not intended to replace the care provided by a licensed healthcare professional, including prescriptions, diagnosis, or treatment.

BlueStar[®]は
糖尿病の自己管理を
助けるもの

BlueStar[®]は
医師の処方、診断、
治療をreplaceする
ためのものではない

FDAで試験的な制度が運用開始: Software Precertification (Pre-Cert) Program

- 現行の医療機器の審査制度は、迅速な進歩と改良を伴うSaMDには、**そぐわない(not-well suited)**
 - ◆ 数カ月～数年単位の改良を伴う医療機器製品と異なり、ソフトウェア製品は、**real-world performance**と**user feedback**に基づき、数週間～数カ月単位で**update**される
- Pre-Certプログラムは、合理化・迅速化された審査体制を構築するためのトライアル
- 2つのレベルで**審査**と**Real-Worldフィードバック**を反復:
SaMD規制のTotal Product Lifecycleアプローチ
 - ◆ Organization-Level Analytics:
企業の体制・質・信頼性の評価
 - ◆ Product-Level Analytics:
製品レベルの評価(審査項目・重さの決定)
- 製品の**safety & effectiveness**の合理的な保証が目的

Developing a Software Precertification Program: A Working Model

v1.0 - January 2019

<https://www.fda.gov/media/119722/download>

製品リスクと
企業レベルで分岐

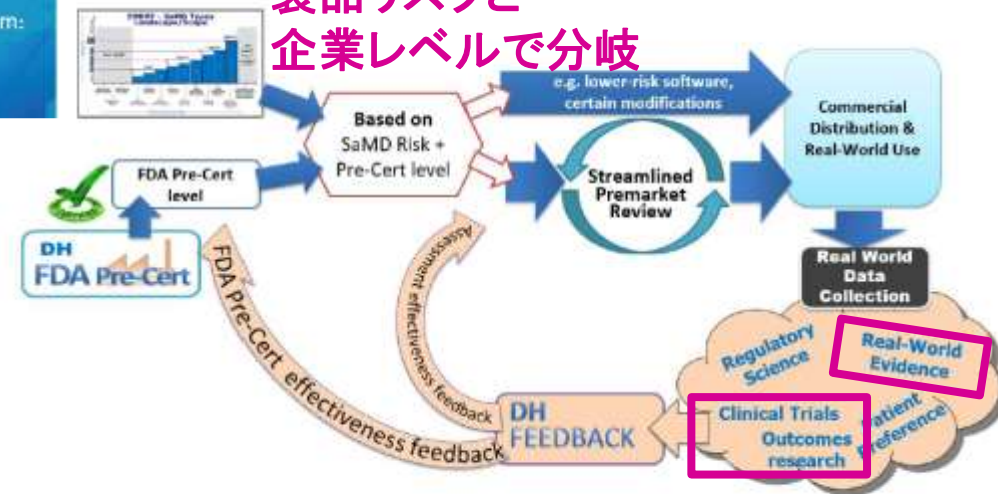


Figure 1. A reimagined approach for the regulation of software

企業レベルと製品レベルで循環

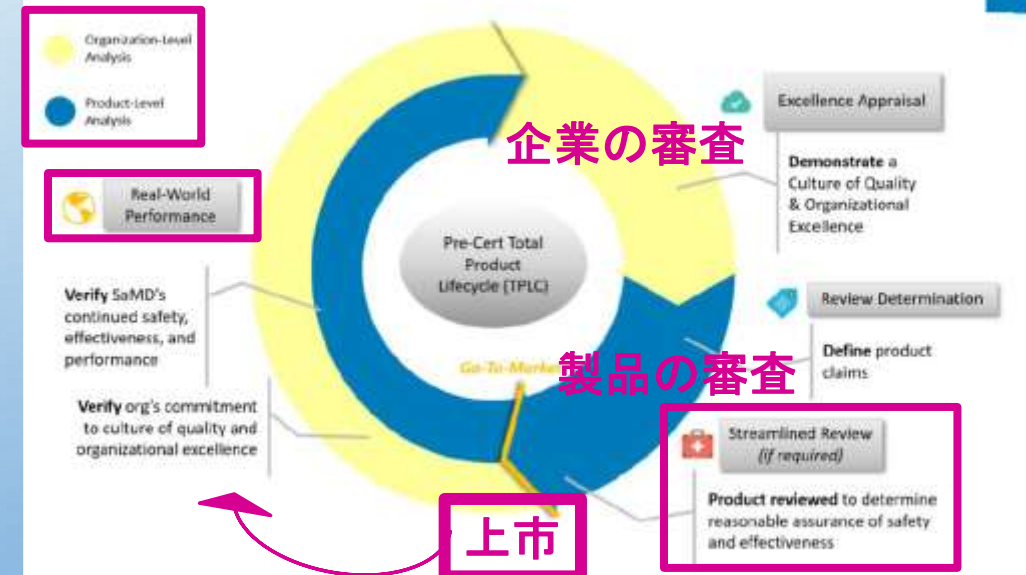


Figure 4. Total Product Lifecycle Approach of the Software Pre-Cert Program

上市前の審査項目

■ Safety & Effectiveness

の保証に必要な審査項目

◆Cybersecurity

◆Clinical Data Analysis & Interpretationなど

✓製品単位のClinical Performanceを示す臨床データの解析と解釈

✓リスクレベルや企業レベルにより審査の深さは異なるとのこと

6.1 Elements necessary for assuring safety and effectiveness in premarket review

As described in Section 3.5 above, FDA proposes that certain elements traditionally reviewed in a premarket submission for a SaMD product can be evaluated at the organization level during the Excellence Appraisal and at the product level during Review Determination. The following table includes the elements that would be reviewed during a Streamlined Review in order to provide a reasonable assurance of safety and effectiveness at the point of market entry.

Table 5. Streamlined Review Elements

Administrative Elements	
Cover letter	
Financial Certification and Disclosure Form	
Truthful and Accuracy Statement	
Product-Specific Elements	
Clinical algorithm	
Clinical Data Analysis and Interpretation	
Cybersecurity product-specific information including threat model	
Declaration of Conformity and Summary Reports for Vertical Standards	
Hazard Analysis (product-specific)	
Instructions for use	
Labeling review	
Regulatory Pathway Specific Items (e.g., 510(k) substantial equivalence comparison)	
Requirements (product-specific)	
Revision history	
SaMD product demo	
Software architecture	
Validation (product performance)	
Elements Leveraged from other components	
Excellence Appraisal Assessment	
Review Determination information (Indications for Use, Device Description, etc.)	
Real-world Performance Plan	

DiGAディレクトリへの「暫定的」追加という制度を実験的に施行 ドイツの治療用アプリに対するFast Track制度

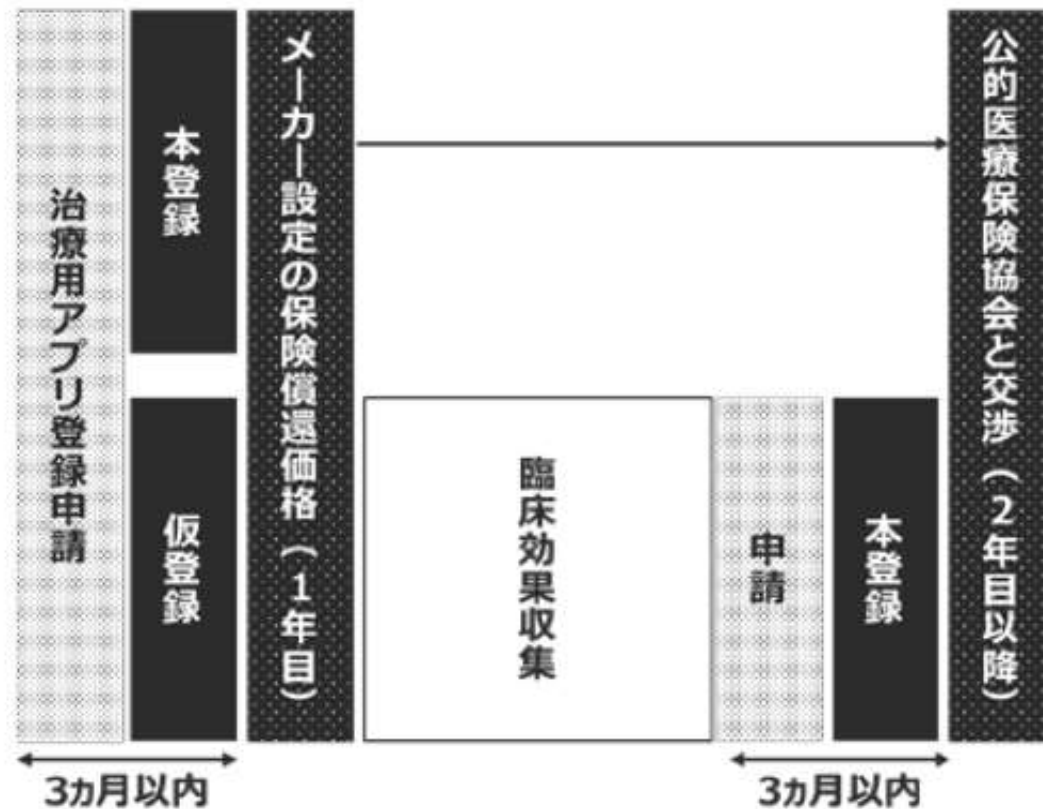


図 6 ドイツの治療用アプリ登録・保険償還制度 ¹⁷

- 登録要件を満たせば、申請から3カ月以内に処方可能、**保険償還価格**も決まる
- 仮登録後、臨床効果に関するRWDを収集、本登録
◆ 対照群との**比較が必要**だが、**RCT**は必須ではない

日本総研.「医療のデジタル化におけるデジタルセラピューティクス(DTx)導入の推進に関する提言」より抜粋

ドイツにおける承認・開発状況

DiGAの実例

DVG施行から1年あまりが経過し、ドイツにおいては同法の下で数多くのアプリが社会実装されました。2021年7月18日現在、ディレクトリに追加されているDiGAは表3のとおりです。一方、日本においては、疾病の治療に用いられるプログラム医療機器として保険償還の対象となっているのはCureApp SCのみです。（なお、CureApp SCはハードウェアであるCOチェッカーとのセットの製品ですので、ドイツにおいてはDiGA対象とはなりません。）

製品名	会社名	適応症	登録日	価格 (Euro)	登録方式
Kalmeda	mynoise GmbH	耳鳴り	2020/9/25	116.97	暫定的追加
velibra	GAIA AG	パニック障害	2020/10/01	476	本追加
somnio	mementor DE GmbH	不眠症	2020/10/22	464	本追加
Vivira	Vivira Health Lab GmbH	腰痛・膝痛・股関節痛	2020/10/22	239.97	暫定的追加
zanadio	aidhere GmbH	肥満	2020/10/22	499.8	暫定的追加
Invirto	Sympatient GmbH	パニック障害	2020/12/03	428.4	暫定的追加
elevida	GAIA AG	多発性硬化症	2020/12/15	43.75	本追加
M-sense Migraine	Newsenselab GmbH	片頭痛	2020/12/16	219.98	暫定的追加

製品名	会社名	適応症	登録日	価格 (Euro)	登録方式
Selfapy's Online Course on Depression	Selfapy GmbH	鬱	2020/12/16	540	暫定的追加
Rehappy	Rehappy GmbH	脳卒中の方のアフターケア	2020/12/29	449	暫定的追加
deprexis	GAIA AG	鬱	2021/2/20	297.5	本追加
Mika	Fosenis GmbH	子宮頸がん、子宮がん、卵巣がん	2021/3/25	419	暫定的追加
Mindable: Panic Disorder and Agoraphobia	Mindable Health GmbH	パニック障害、広場恐怖症	2021/04/29	576	暫定的追加
CANKADO PRO-React Onco	CANKADO Service GmbH	乳がん	2021/05/03	499.8	暫定的追加
vorvida	GAIA AG	アルコール依存症等	2021/05/06	476	本追加
Selfapy's Online Course for panic disorder	Selfapy GmbH	パニック障害	2021/6/19	540	暫定的追加

運用開始約1年間で

仮登録14品目
本登録5品目

製品名	会社名	適応症	登録日	価格 (Euro)	登録方式
Selfapy's Online Course on Generalized Anxiety Disorder	Selfapy GmbH	全般性不安障害	2021/6/19	540	暫定的追加
Non-smoking Hero app	Non-smoking Hero app	禁煙	2021/7/3	239	暫定的追加
Esysta - Digital diabetes management	Emperra GmbH	1型及び2型糖尿病	2021/7/4	249.86	暫定的追加

表3：DiGAディレクトリに登録されている製品

本邦では、どんなエビデンスがあればよいか？

- ランダム化比較試験
- アプリ群 vs. シャムアプリ群
- 確立された臨床エンドポイントや代替エンドポイント
- 臨床的に意味のある差



綺麗な形のエビデンスを示してくれれば皆納得する

厳密な有効性評価はどこまで必要か？
何かうまいやり方はないか？

A blue-toned financial candlestick chart with a grid background. A white curved line, likely a moving average, trends downwards from the top left. Several numerical values are highlighted in white boxes: '104.19' near the top, '86.72' in the middle left, and '72.48' at the bottom. A label '61.6%: 99.19' is positioned near the top right. The chart is decorated with translucent blue bubbles of various sizes.

統計的な アプローチ **PRECISION STATISTICS**

革新的な試験デザインによる効率的なデータ収集

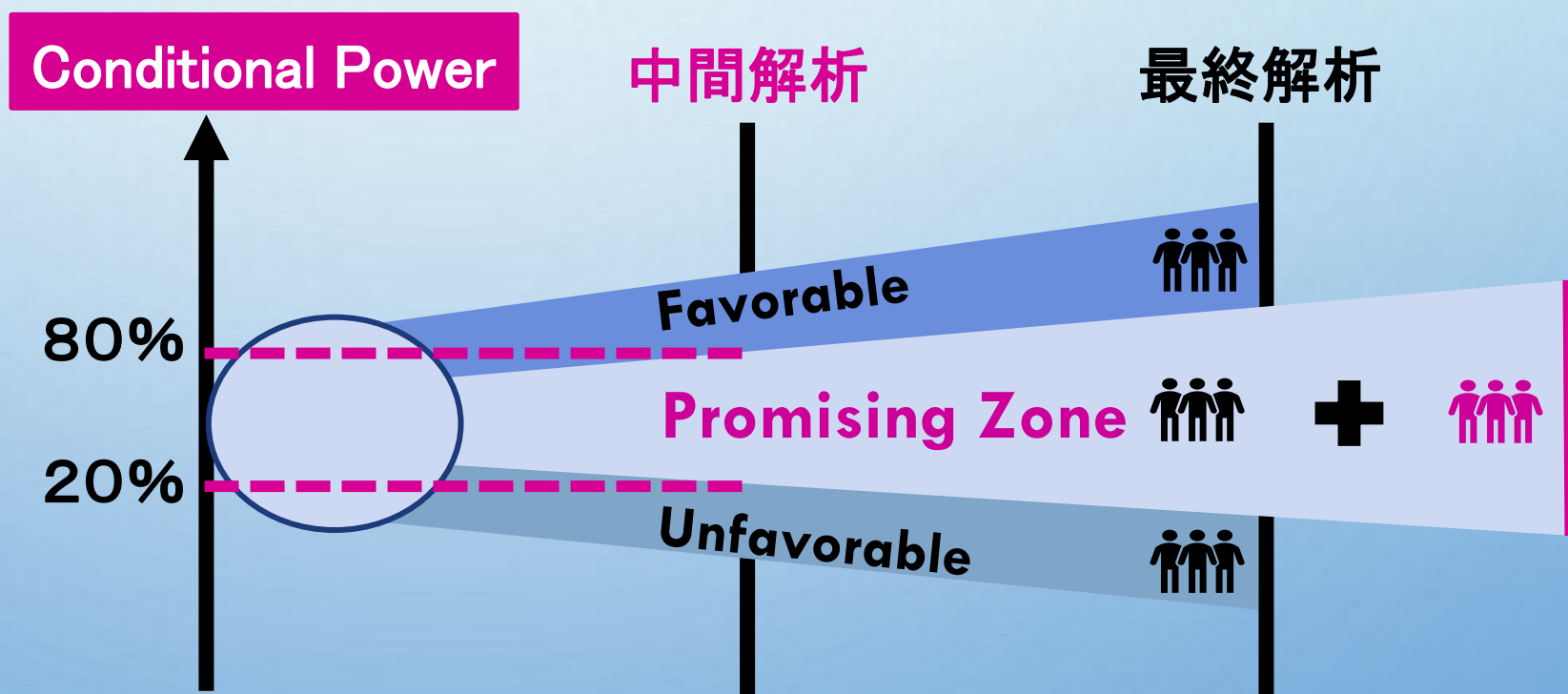
- アダプティブ・デザインを活用した試験
- Real World Dataを活用したHybrid Control Design
- 有用性の評価に適したエンドポイントの開発
- アプリに対する反応の良い集団を特定するデザイン

Adaptive Designの例

■ Sample Size Re-estimation (SSR)

◆ 症例数の見積もりが甘かった場合、**中間解析**で見直し

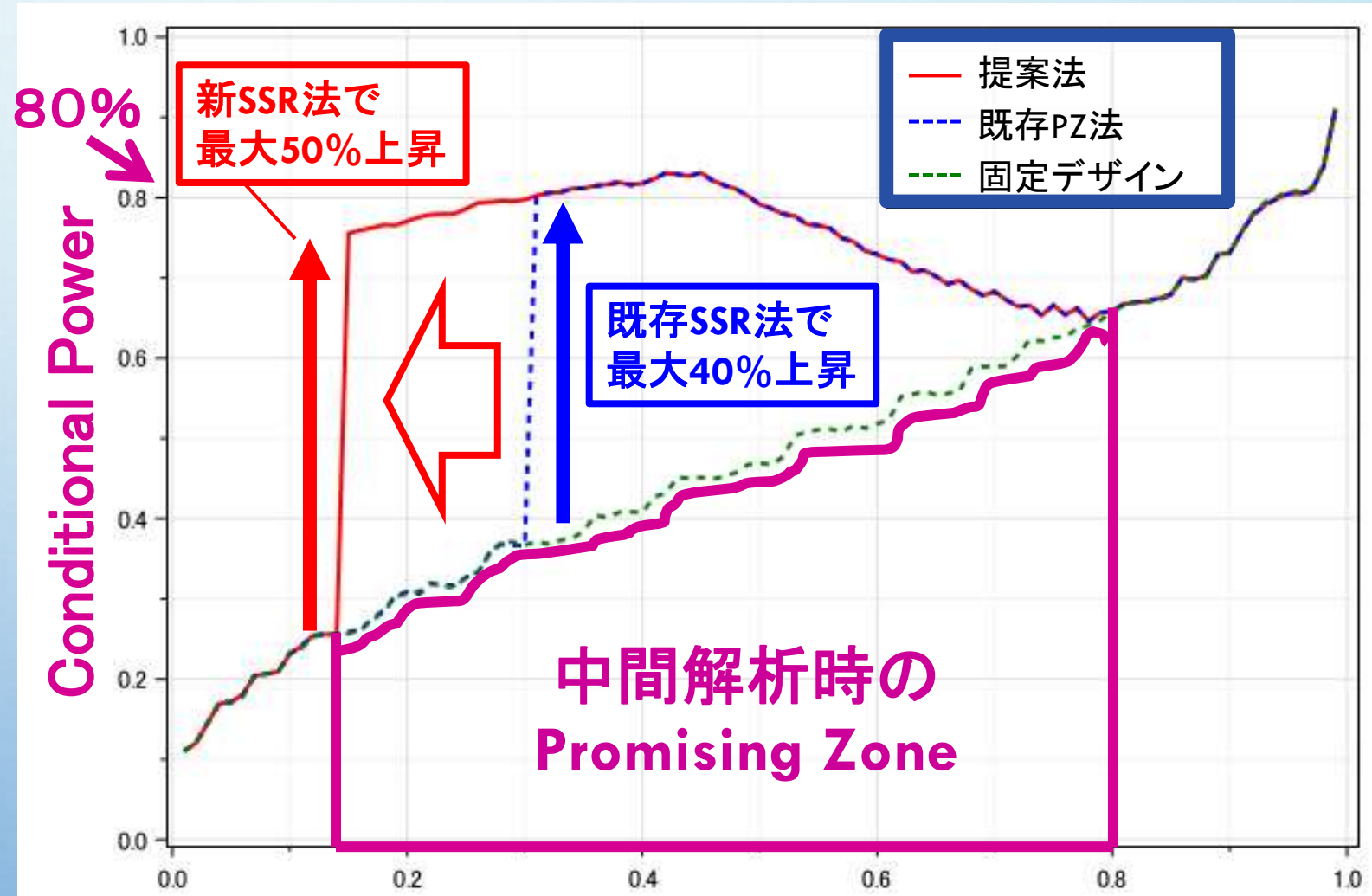
◆ **有望な結果であれば、**
試験をやり直す
のではなく、
拡大して継続



SSRの適用例: Trial Simulation

※ 東大AROの高嶋郁海さんの
修士論文発表スライドを改編

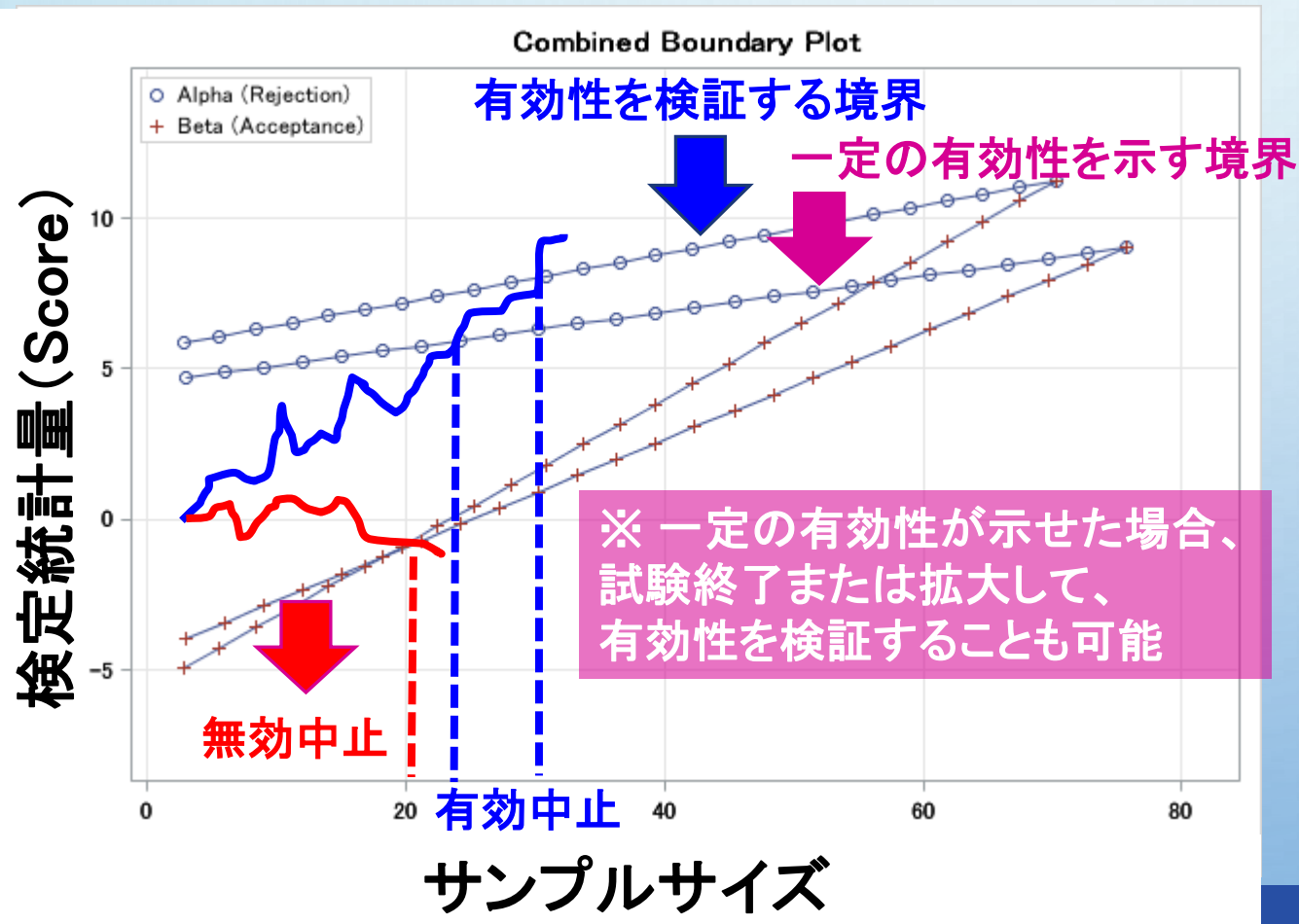
- **Conditional Power**
 - ◆ 中間解析時に判定した試験の成功確率
- **失敗に終わる試験を確度高く救える**
 - ◆ Nmaxは300例
 - ◆ Ninitialは150例
 - ◆ Nmeanは180例



小規模なStudyで有効性を推定、可能ならば検証

■ Sequential Study based on **Double Boundaries** Approach

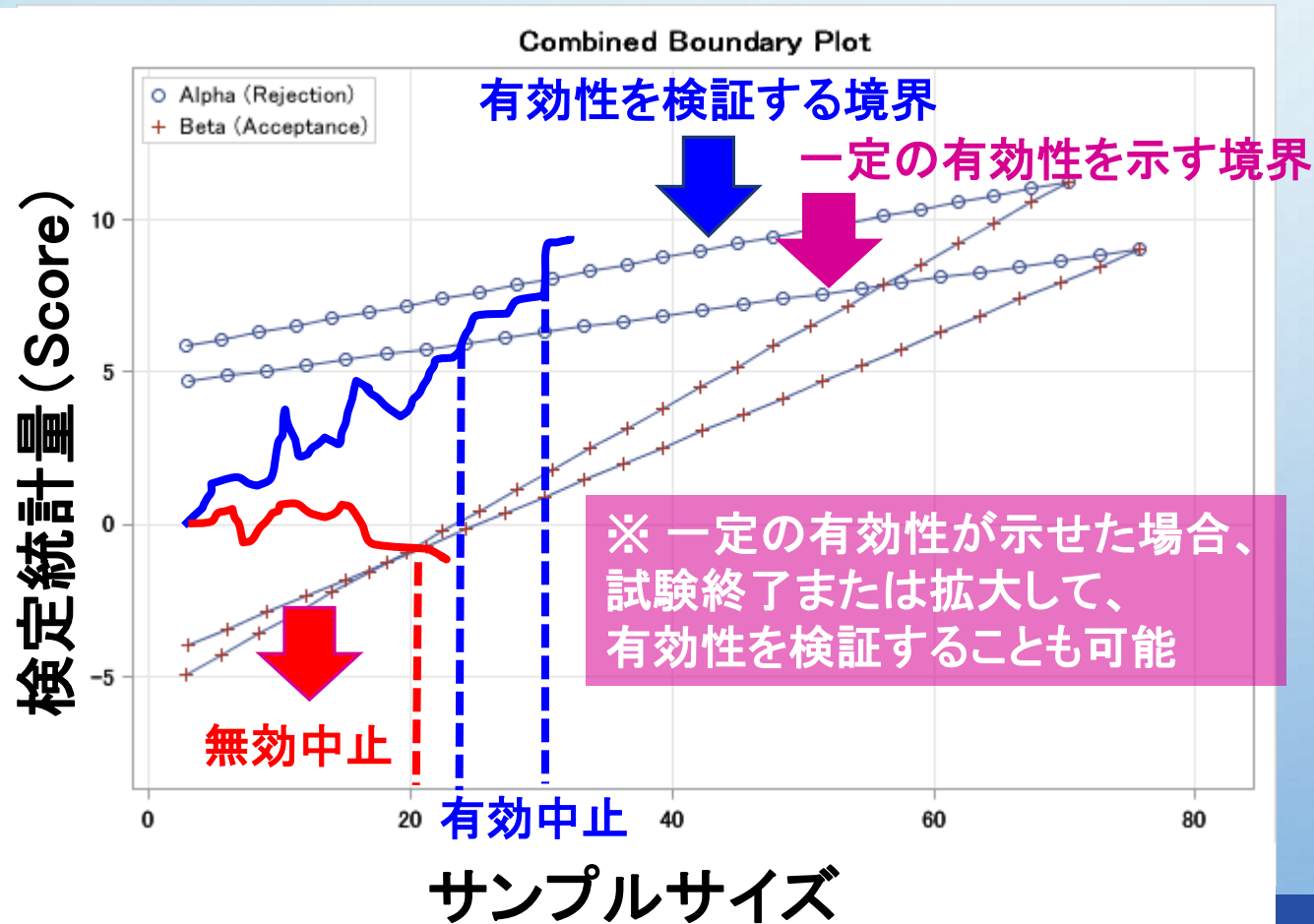
◆ 検定統計量を逐次モニタリングし、境界を越えたら中止



小規模なStudyで有効性を推定、可能ならば検証

■ Sequential Study based on **Double Boundaries** Approach

◆ 検定統計量を逐次モニタリングし、境界を越えたら中止

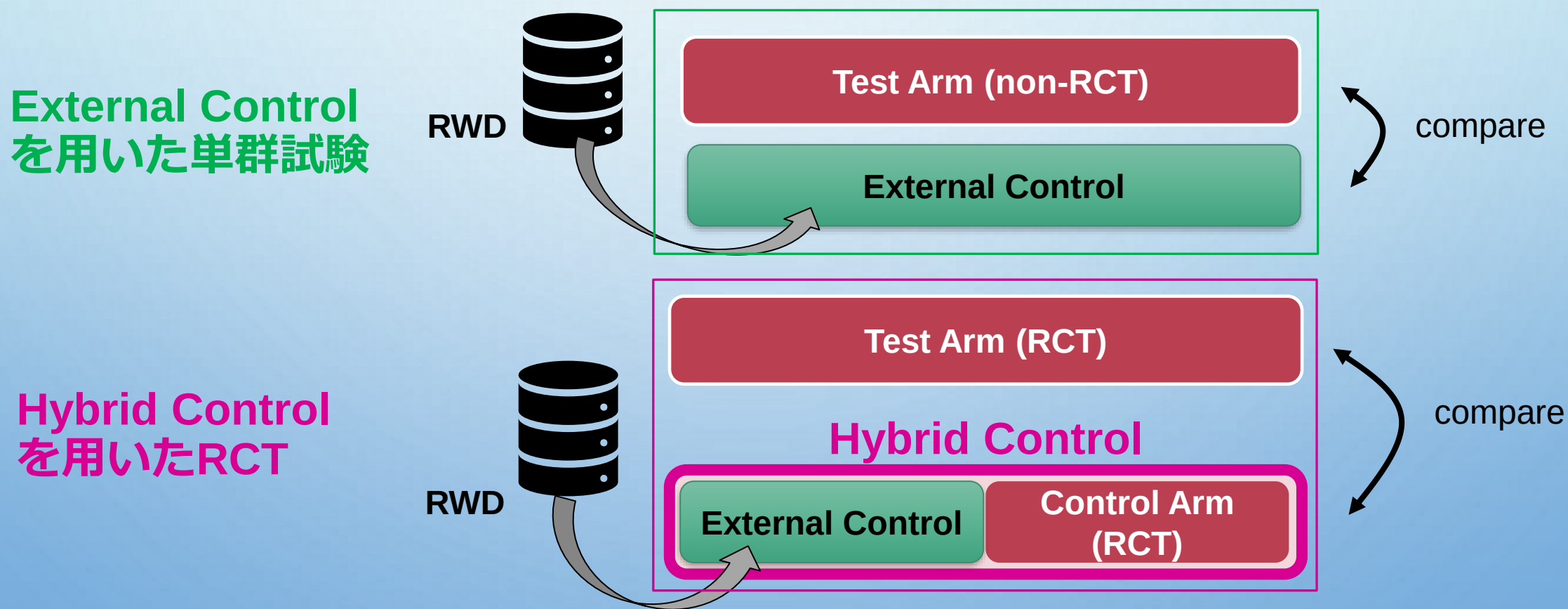


- 通常のデザインだと50例必要(検出力80%)
- このアプローチだと**平均34例**で有効性を検証
- 無効な場合は**平均26例**で**早期中止**できる

ランダム化への同意が難しいケースなど

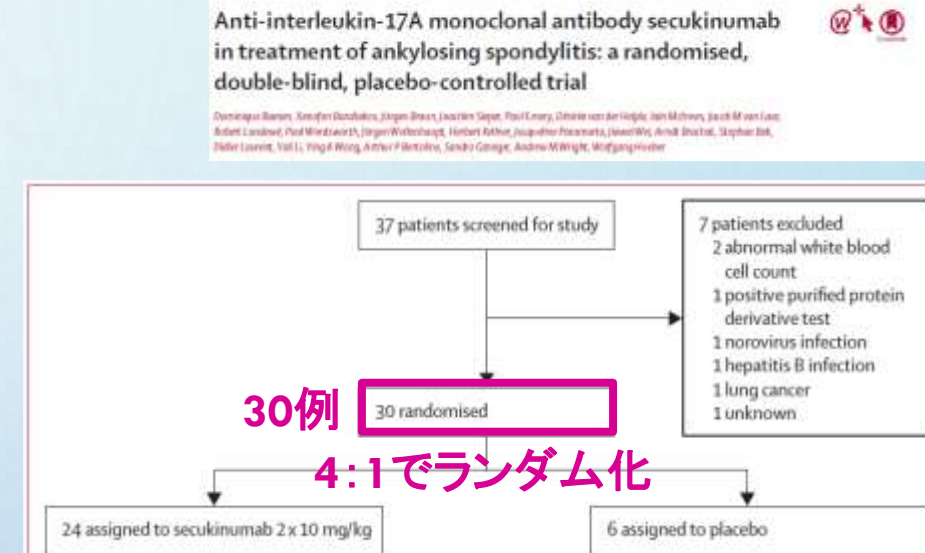
■ External Controlとの比較 vs. Hybrid Controlとの比較

◆ランダム化の良さを生かし、**Type I error (bias)**をある程度制御したい



Lancet 2013; 382: 1705–13

-
- Figure 1 is a sagittal MRI scan of the spine. A white box highlights a vertebral fracture at the T12 level. A line points from this box to a larger inset image on the right, which provides a magnified view of the fracture site. The spine is shown in profile, with the vertebrae and intervertebral discs clearly visible. The fracture is located in the vertebral body of the T12 vertebra.



	Responders, n (%)	Response rate*	Difference vs placebo†	95% credibility interval†	Probability
Secukinumab‡§	14 (60.9%)	59.2%	34.7%	11.5–56.4%	99.8%
Placebo	1 (16.7%)	24.5%	..	..	..

ASAS=Assessment of SpondyloArthritis international Society criteria. *Means from the posterior beta (0.5 + x, 1 + n - x) distribution for secukinumab and beta (11 + x, 32 + n - x) distribution for placebo, where x represents the number of responders and n - x represents the number of non-responders in the corresponding treatment group. †Difference in response rates simulated from the posterior probability distributions of secukinumab and placebo. ‡Secukinumab: 2 × 10 mg/kg. §The efficacy dataset included only 23 of 24 patients in the secukinumab group, since one patient was excluded due to a dosing error.

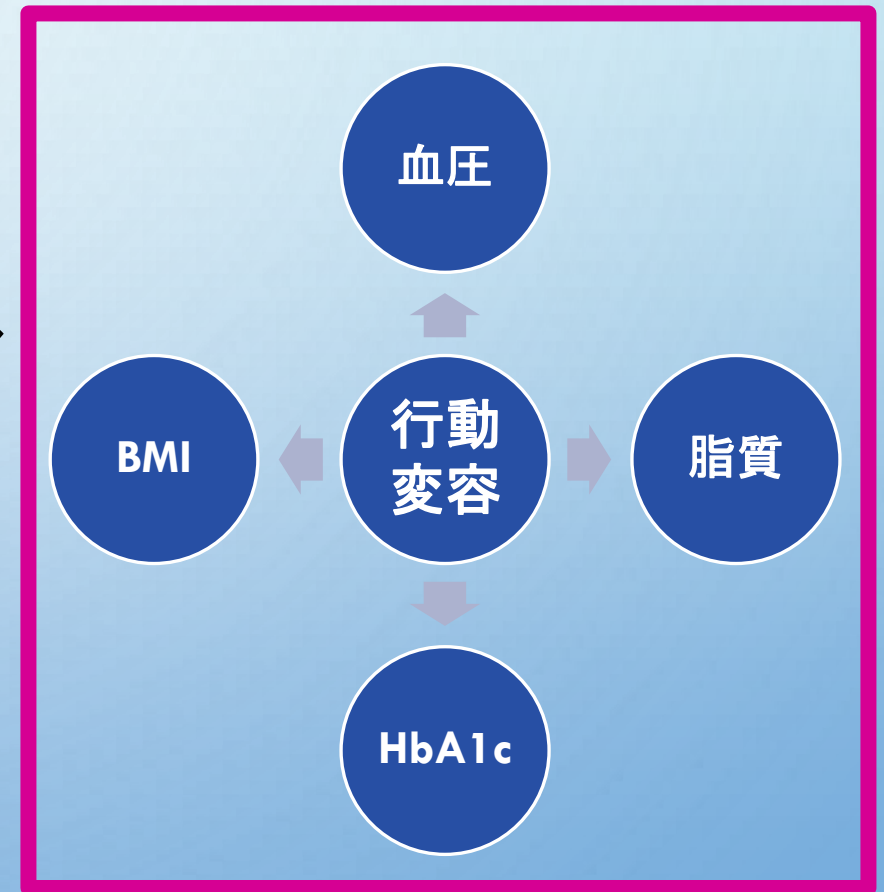
東大シンポジウム「行動変容プログラム/アプリの開発と評価」

エンドポイントの設定に問題が生じる例

■ 生活習慣病に対するアプローチ

- ◆BlueStarのように疾患特異的なアプリも良いが、リスク因子は単一ではない
- ◆**複数のリスク因子**を持つ幅広い患者層を、一括管理できるアプリも有用
- ◆薬のように作用点が1つではないため、
 - ✓**個々のエンドポイントに対する短期的効果は必ずしも小さくなく、有用性を示すのは困難なケースもある**

単独の数値だけを改善するよりも、
複数のリスク要因を同時に管理！



リスクスコアで1次元の尺度へ

- 仮にリスクスコアをモニタリングし、行動変容させるアプリがあると、
 - ◆結果的に**リスクスコアが低下**すれば、有用性を示すには十分と考えられる
 - ◆複数の要因が改善するのであれば、**統計的なPowerも大きく改善する**
- リスクスコアの作成方法
 - ◆**JPHC Study等の疫学研究をベース**

ORIGINAL CONTRIBUTIONS

Development of a Point-based Prediction Model for the Incidence of Total Stroke

Japan Public Health Center Study

Hiroshi Yatsuoka, MD, PhD, Hiroyasu Iso, MD, MPH, PhD, Kazumasa Yamagishi, MD, PhD, Isao Saito, MD, PhD, Kazuo Suzuki, MD, Norie Inoue, MD, PhD, and Shoichiro Tsugane, MD, PhD

Table 2. Points Assigned to Stroke Risk Factor Categories Japan Public Health Center Study Cohort II, 1993 to 2007

Risk Factor	Points
Age group, y	
40–44	0
45–49	5
50–54	6
55–59	12
60–64	16
65–69	19
Sex	
Men	6
Women	0
Current smoking	
Men	4
Women	8
Body mass index, kg/m ²	
<25	0
25 to <30	2
30–	3
Blood pressure, mm Hg	
No antihypertensive medication	
<120/80	0
120–129/80–84	3
130–139/85–89	6
140–159/90–99	8
160–179/100–109	11
>180–/110	13
On antihypertensive medication	
<120/80	10
120–129/80–84	10
130–139/85–89	10
140–159/90–99	11
160–179/100–109	11
>180–/110	15
Diabetes mellitus	
Yes	7
No	0

脳卒中の
ハザード比

ポイント化

各因子の
ポイントの
合計点

リスク
スコア

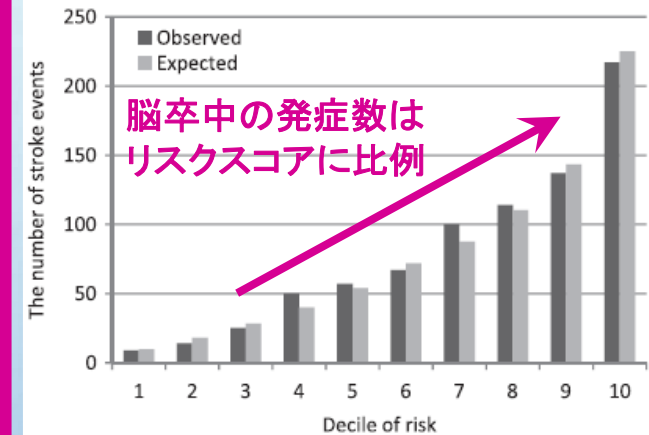


Figure 2. Calibration of the prediction models by decile of risk. Observed (solid bar) and expected (gray bar) numbers of strokes according to deciles of 10-year risk showed good agreement in

Table 3. Predicted 10-Year Risk of Stroke by Risk Score and Corresponding Vascular Age, Japan Public Health Center Study Cohort II, 1993 to 2007

Score	Predicted Risk (%)	Vascular Age, y	
		Men	Women
≤10	<1	42	47
11–17	1 to <2	53	60
18–22	2 to <3	59	67
23–25	3 to <4	64	72
26–27	4 to <5	67	76
28–29	5 to <6	70	80
30	6 to <7	73	83
31–32	7 to <8	75	85
33	8 to <9	77	>90
34	9 to <10	79	>90
35–36	10 to <12	82	>90
37–39	12 to <15	85	>90
40–42	15 to <20	>90	>90
43–	20–	>90	>90

治療用アプリに**反応の良い集団**を探索することも重要:

マイクロ・ランダム化

- 観察研究とランダム化を組み合わせたアプローチ
- 治療用アプリによる介入のタイミングや条件の探索・最適化

↑
**Phase II/III試験もしくは
上市後のReal World Performanceの評価**
↓

ランダム化治療中止実験

- 治療用アプリに反応した集団に対し、治療用アプリによる介入の継続の有無をランダム化

Circ Cardiovasc Qual Outcomes. 2021;14:

Circulation: Cardiovascular Quality and Outcomes

NOVEL STATISTICAL METHODS

Microrandomized Trial Design for Evaluating Just-in-Time Adaptive Interventions Through Mobile Health Technologies for Cardiovascular Disease

ABSTRACT: Smartphone and wearable device use is rising broadly and can be leveraged for chronic disease management. Just-in-time adaptive interventions promise to deliver personalized, dynamic interventions directly to patients through use of push notifications from mobile devices. Although just-in-time adaptive interventions are a powerful tool for shaping health behavior, their application to cardiovascular disease management has been limited as they can be challenging to design. Herein, we provide a general overview and conceptual framework for microrandomized trials, a novel experimental study design that can be used to optimize just-in-time adaptive interventions. Microrandomized trials leverage mobile devices to sequentially randomize participants to types or levels of an intervention to determine the effectiveness of an intervention and time-varying moderators of those effects. Microrandomized trials are an efficient study design that can be used to determine which intervention components to include in just-in-time adaptive interventions and to optimize their decision rules while maintaining the strength of causal inference associated with traditional randomized controlled trials.

Jessica R. Golbus¹, MD, MS
Walter Dempsey, PhD
Elizabeth A. Jackson, MD, MPH
Brahmaje K. Nallamothu², MD, MPH
Predrag Klasnja, PhD

ご清聴ありがとうございました



BACK UP

有効性の評価 vs. 有用性の評価

■ 行動変容アプリの位置づけによる

- ◆ 治療自体を担うのであれば、患者の臨床的ベネフィットを直接表す臨床エンドポイントまたは期待できる代替エンドポイントにより、有効性を評価すべきと考えられる
- ◆ 治療の補助や患者の管理・行動変容を担うのであれば、患者の臨床的ベネフィットだけでなく、位置づけに見合った評価方法を検討する必要がある

治療薬と治療用アプリの違いは何か

- リスク・ベネフィットバランスが異なる場合が多い
- 薬価よりもアプリの診療報酬の方がはるかに安い
- 安くてうまいのであれば、どんどん積極的に導入すればよい
 - ◆ 使い勝手が悪い、役に立たないのであれば、自然淘汰される
 - ◆ 改良を重ね、市場で生き残ったSaMDには、それなりの価値がある

どの程度の厳密な綺麗なエビデンスが必要なのか？

- 綺麗なエビデンスは文句が付きにくく、審査をする側にとっては大変都合がよいのは確かです
 - ◆ エビデンスは国のためではなく、国民のために構築していくべきであるのも厳然たる事実
- 個人的には、一定の保証をするに足る証拠の提示により、お墨付きを与えるという役割が大きいのではないかと？
 - ◆ 「デジタルトクホ」のようなイメージ
- 比較試験といえで、盲検化が難しい
 - ◆ 試験へ参加した集団は、コンプライアンスが現場と異なるかもしれない
 - ◆ ちゃんと使い続けられるかどうかは、アプリの使用感や効果の実感や個人の意思による
 - ◆ 役に立たない製品は、自然と淘汰される可能性があり、市場に出てからいかにアップデートしていくか
- 開発コストが医薬品と比較すると安価とはいえ、新規参入企業にとっては、回収は容易ではない。わざわざ医療用として開発し、医師に使用してもらうためには、うまい仕組みにしないと、発展しないのでは？と思う。
 - ◆ 大きな企業にとっては、ハードルが高めの方が、競争の激化を防げるため、良い面もあるのかもしれない
 - ◆ 周回遅れになっていない今だからこそ、議論する価値があるのではないかと。SaMDラグや審査のやり方や相場観が定着してしまってから、なかなか変えられないし

- 諸外国の状況を見ると、イギリスでは医療におけるデジタル分野の開発に関するガイダンス等の提示と更新を進めている。ドイツでは、治療用アプリは申請後 3 カ月以内に薬事承認・保険償還され、有効性のエビデンスが十分でないアプリは仮償還として収載し、1 年後に再度保険償還に挑戦できる制度を施行している。そして、中国においては 2020 年 11 月に、中国初の治療用アプリが薬事承認され、国内の注目が集まっている。
- 中でも、ドイツは健康増進だけでなく、開発促進や薬事承認・保険償還早期化の観点からも政策を進めており、他国でも政策の深掘りが進むと想定される。河