

第3回モニタリング2.0検討会シンポジウム「SDVの効率化を考える」

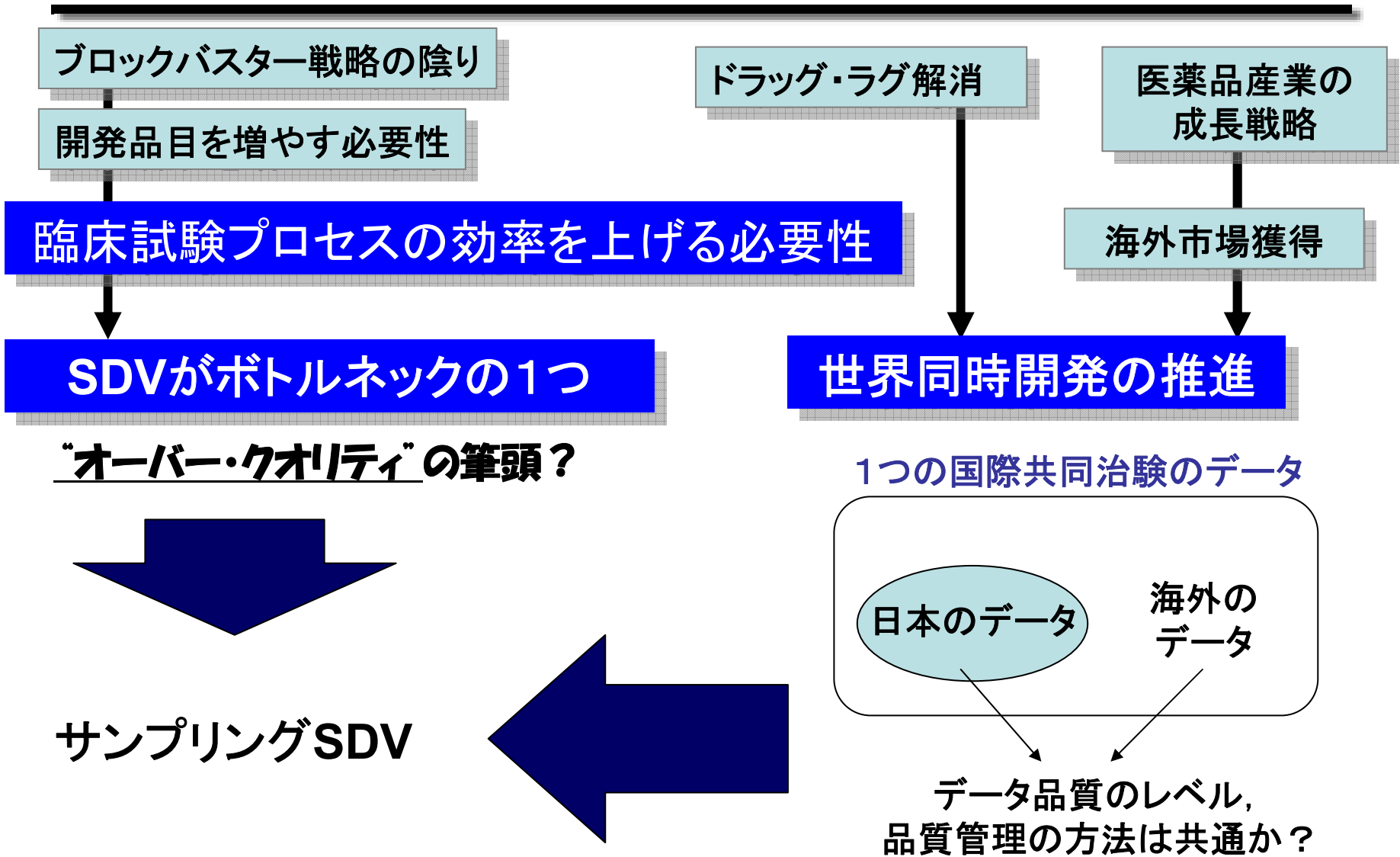
サンプリングSDVへの挑戦

ファイザー株式会社

臨床統計部・統計コンサルティンググループ

小宮山 靖

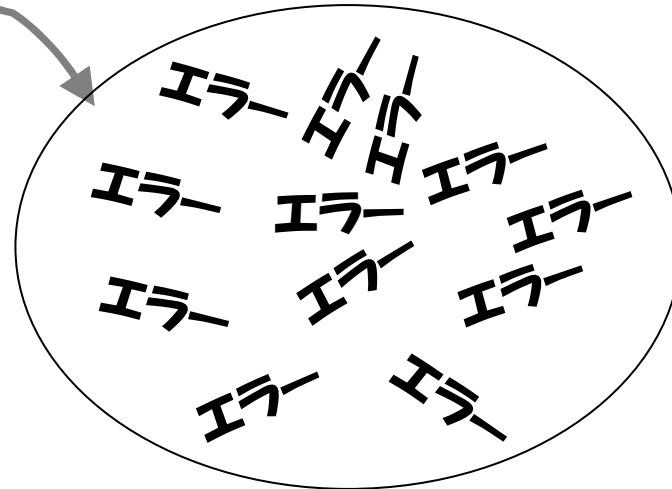
なぜサンプリングSDV？



●● サンプルングで仕事を間引こう！



ムダなデータ、
ムダなプロセスで
肥大化したお仕事のかたまり
エラーもたくさん含まれる



こんなことをやっちゃ
いけません

見逃しも
多くなる



データ品質が悪い場合には

CRFに記録された治験データ(施設毎)

エラー エラー エラー エラー エラー エラー
エラー エラー エラー エラー エラー エラー
エラー エラー エラー エラー エラー エラー
エラー エラー エラー エラー エラー エラー
エラー エラー エラー エラー エラー エラー

エラー
エラー
エラー

サンプルの中にエラーが
たくさん見つければ、
チェックしていない部分にも
たくさんあるはず

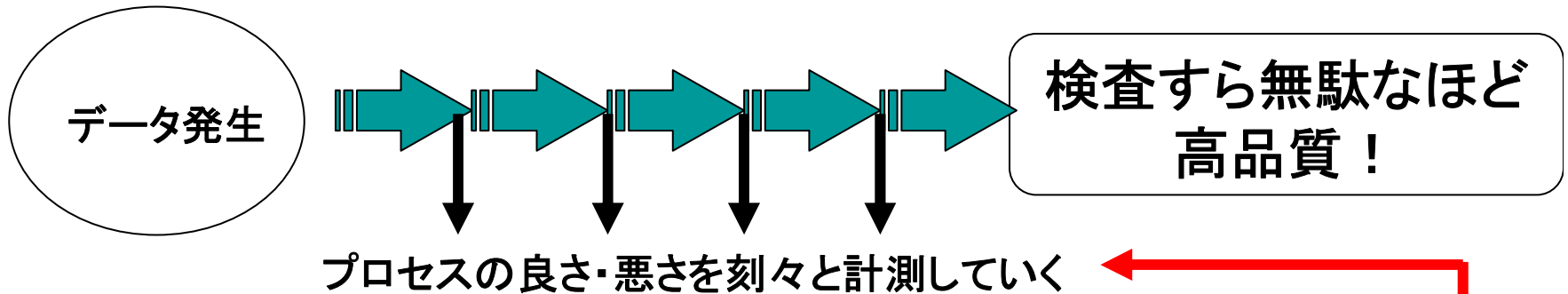
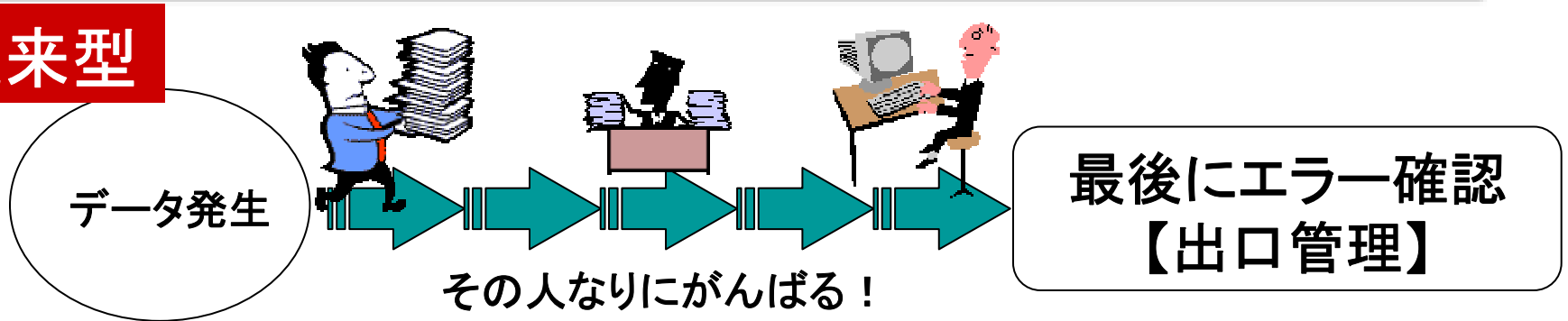
見つかったエラーを修正すればよいのではなく、
そもそもデータ収集プロセスに問題があったということ

●● 根本的な間違いは？

- 「肥大化して、こんがらかったお仕事」に手をつけずに、品質確認だけ手を抜こうとしている
 - ムダなデータ収集をしていないか？
 - ✓ ムダの温床になりやすいのは、病歴、併用治療
 - スポンサー会社や試験に特有な煩雑さはないか？
 - 誤りが起きにくいプロセスになっているか？
 - 試験の進行中にプロセスを見直す努力はしているか？
- 「データが全部集まった後で、最終的に品質が高ければよい(怒られない?)」という稚拙な品質管理の考え方
 - 日本古来のBook型CRFを最後にまとめて回収していた頃の習慣？

品質確認：従来型 vs プロセス管理

従来型



問題があったら、プロセスに悪影響を与える
見逃せない原因 (Assignable Cause) があるはず

その原因を潰していけば、プロセスはますます安定し
品質を作りこむ力 (Process Capability) は向上する！

プロセス管理

品質の作り込み

施設に「品質の作り込み」を要求する前に、やるべきことがある

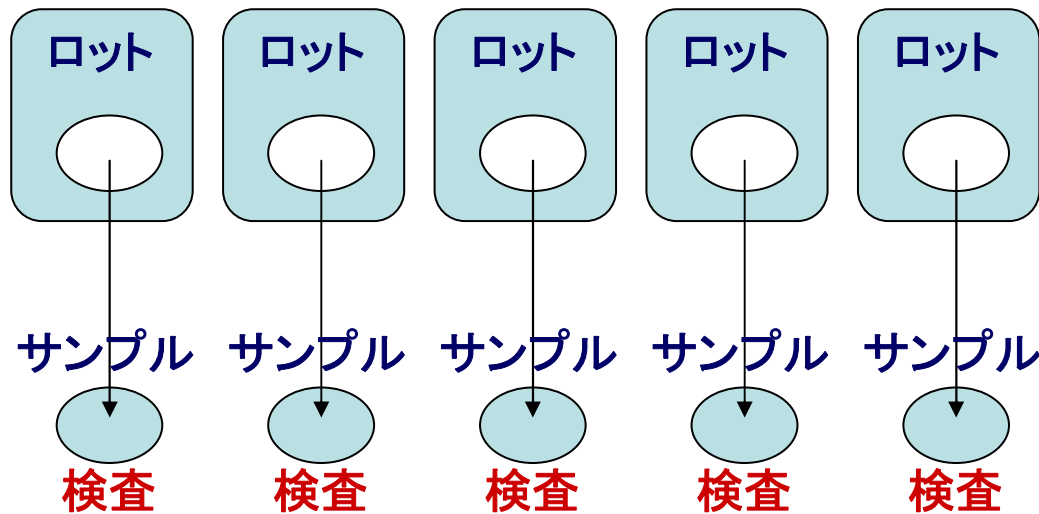
- 簡潔であることを大事にすること(余計なデータを収集しない)
- プロトコールやCRF、CRFの記入手引きなどのコミュニケーションツールの充実(曖昧さの排除)
- 各施設での原資料からCRFまでのデータの流れを把握して、エラーが起きにくい「仕組み作り」に協力すること(Worksheet、カルテに貼るシールなど)
- トレーニングの充実
- 電子化、自動化の推進(人が介入するステップを減らす)

業界全体で取り組むべきこともある

- 標準化
 - 考え方の共通化 ← 「安全性データの収集と報告に関する提言」
 - データ標準の利用 ← CDISC推進

GCDMP翻訳
「安全性データの収集と報告に関する提言」
CDISC推進

何に対してサンプリングを行うか



・・・ ロットは、同一の条件で
(同一の時期に)発生した
データの集まり

サンプリングで品質確認を行い、問題があったときに、
プロセス改善を行う対象は何かを考える

施設？ モニター？ 試験全体？

● ● ● どのようなサンプリングがあるか？

● データ単位

➤ ロットの全データからのランダムサンプリング

✓ データベースに格納されたデータのランダムサンプリング
→ 専用のソフトウェアを設える必要がある(高価)

✓ 手動でランダムサンプリング

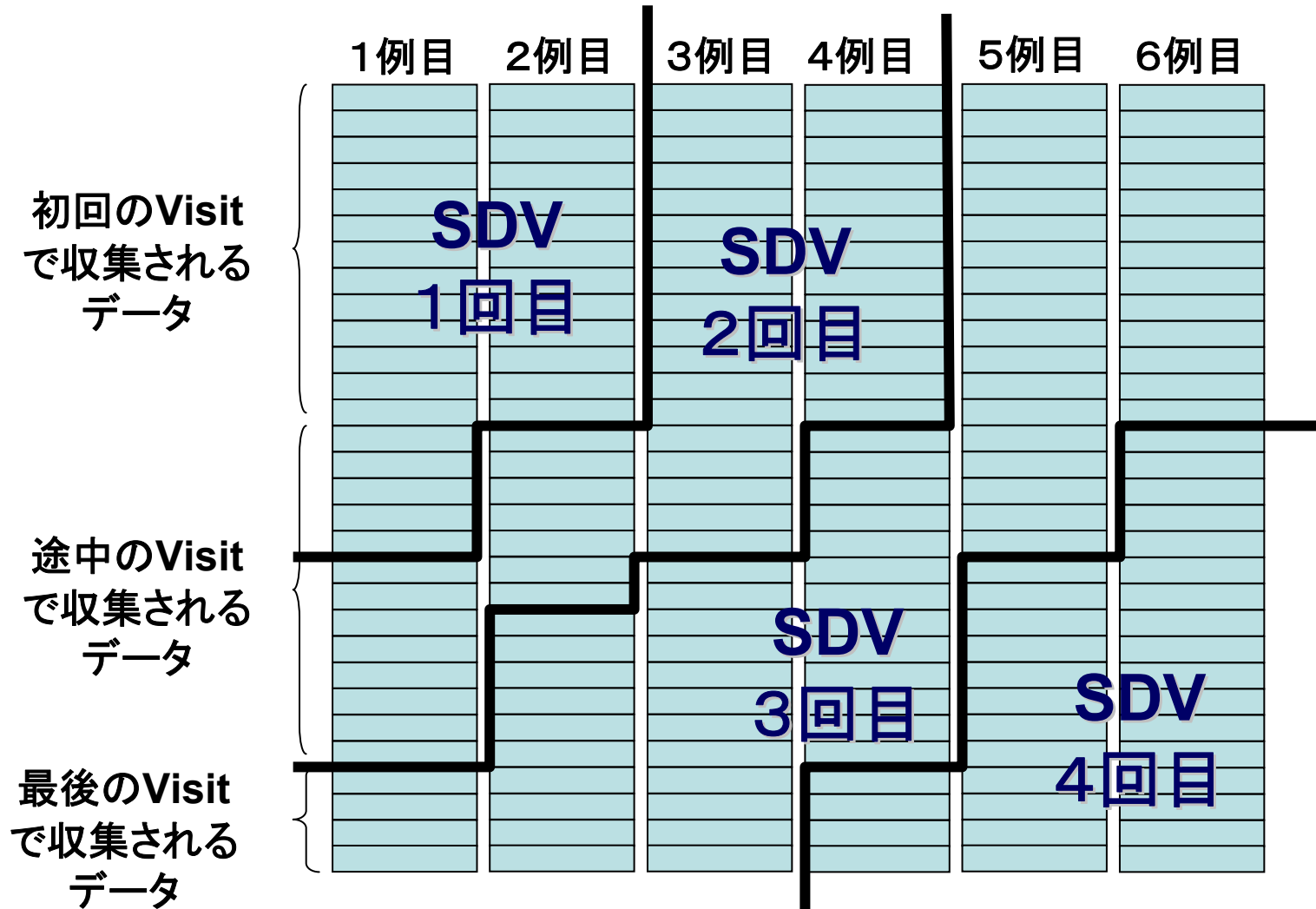
→ ExcelのRAND関数(安価, 手間がかかる)

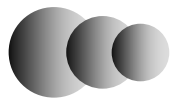
● 症例単位

➤ 各施設の1,3,5,10,15,20...例目は全データをチェック

➤ 全データをチェックする症例をランダムに選ぶ

●● 逐次的にSDVを行うイメージ





手動でランダムサンプリング???

症例

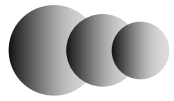
データ項目

	A	B	C	D	E	F	G	H
1								
2						0.332894		
3						0.4503		
4						0.4870204		
5						0.9319615		
6						0.294472		
7					0.696783	0.4094669		
8					0.50032	0.6047835		
9					0.9060263	0.308738		
10					0.6808426	0.5394491		
11					0.7068902	0.168648		
12					0.7636913	0.119259		
13					0.4791546			
14					0.250144			
15					0.5289569			
16					0.5609878			
17					0.4337813			
18					0.321258			
19					0.4820939			
20				0.6706969	0.384477			
21				0.4809724	0.228189			
22				0.8462499	0.365372			
23				0.265093	0.8711947			
24	0.343178	0.5766589	0.9080897	0.357092				
25	0.4153312	0.4866349	0.064726	0.535161				
26	0.8693953	0.4800497	0.9655234	0.278573				
27	0.4045688	0.347233	0.8248206	0.063669				
28								

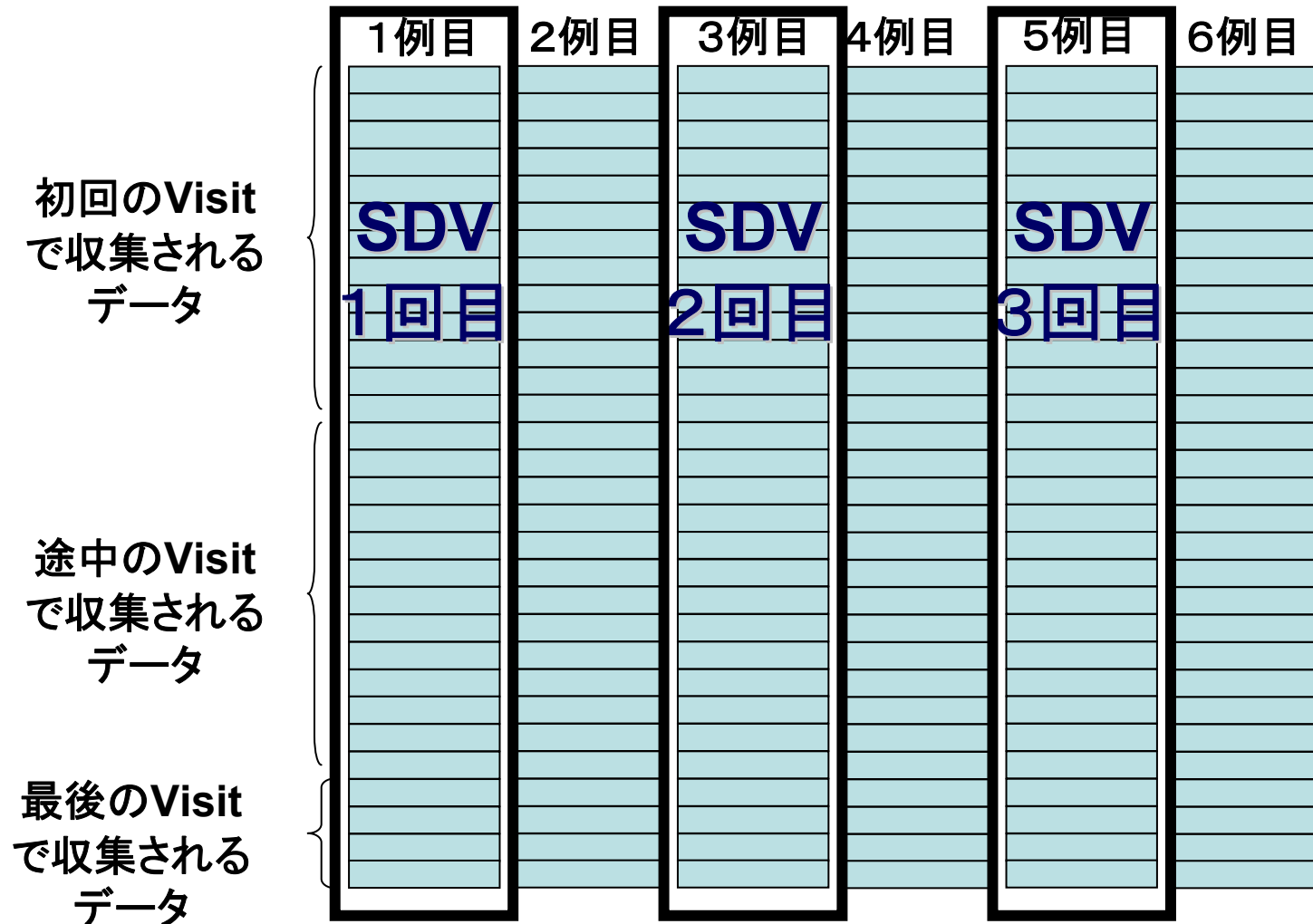
たとえば、
こんなWorksheetを準備。

SDVの対象であるデータ
にExcelのRAND関数で
一様乱数を与える。
昇順に順位付けをして、
上位xx個のデータを
「条件付き書式」機能を使
って、サンプルとして指定

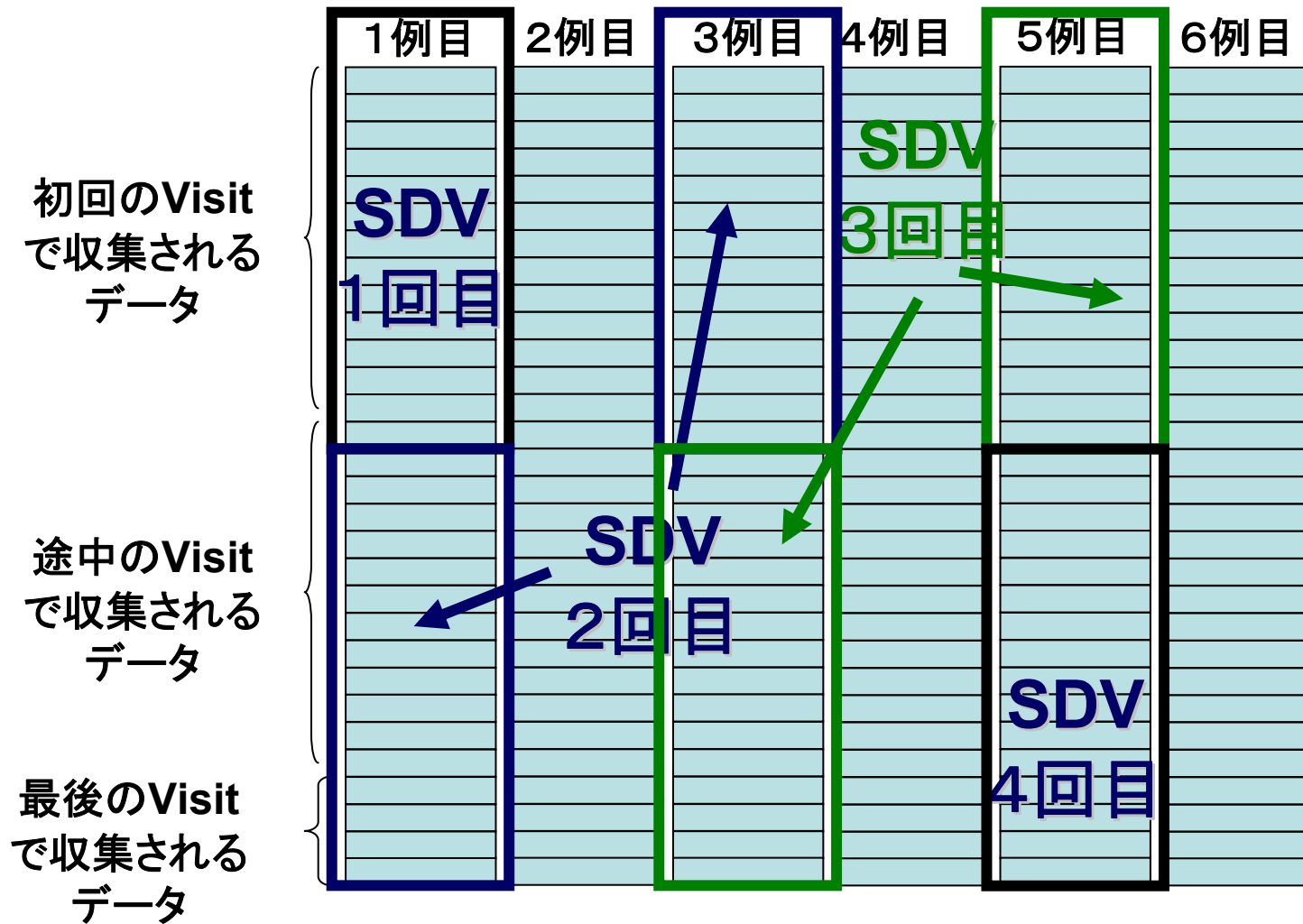
SDV前に、ちょっと
手間はかかるが、
コストはタダ同然

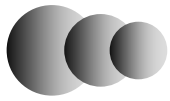


逐次的にSDVを行うイメージ



●● 逐次的にSDVを行うイメージ





全データを均一な基準でチェックするか？ それともデータの重要性でメリハリを付けるか？

● Critical なデータ

➤ SAE関連のデータ

- ✓ 集団での解析がしづらい、
かつ安全対策上とても重要

➤ 同意関係のデータ

- ✓ 治験成立の大前提！

● Critical である「場合」があるデータ

➤ 主要評価項目

- ✓ 主解析のp値が5%近辺で、少数例の結果が全体の結果に影響を及ぼす
（「適切な」サイズの治験なら常に当てはまる？）

➤ 有害事象データ

- ✓ 発現頻度がそれほど高くないAE
（事前に指定できる高頻度のAEはNon-criticalでは？）

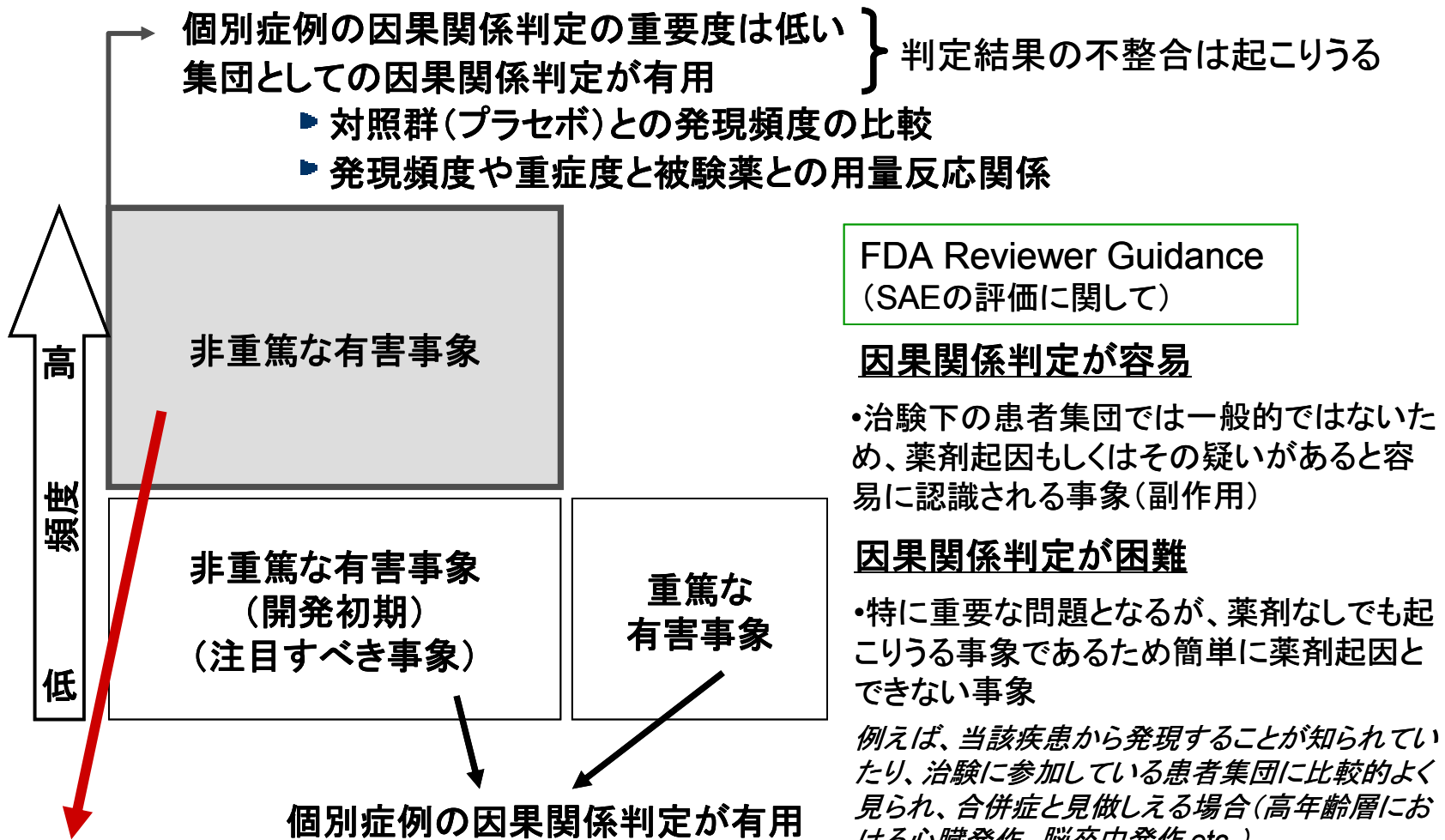
● Non-critical なデータ

➤ 上記以外のデータ

- ✓ SDVに時間がかかっている病歴や併用薬は大抵non-critical

全データを均一な基準で
チェックするのは、
かなり品質が高い状態じゃないと
危険かもしれない！

Criticalな安全性データ？



かなりの部分のAEは事前に特定できるでしょう。
そのAEだけサンプリングの対象にするという手はある。

●● Very Frequently Asked Question

- データ全体のどのくらいの割合をチェックすればいいんですか？
- 合格不合格の判定基準は、どのように設定すればいいんですか？

一概には答えられません。

真面目に統計的抜取検査をやるなら、JISハンドブック57「品質管理」に掲載されているJIS Z 9015 (British Standards BS 6001, ISO 2859などと等価)。これには、サンプルサイズと合否判定の基準の数表が添付されています。

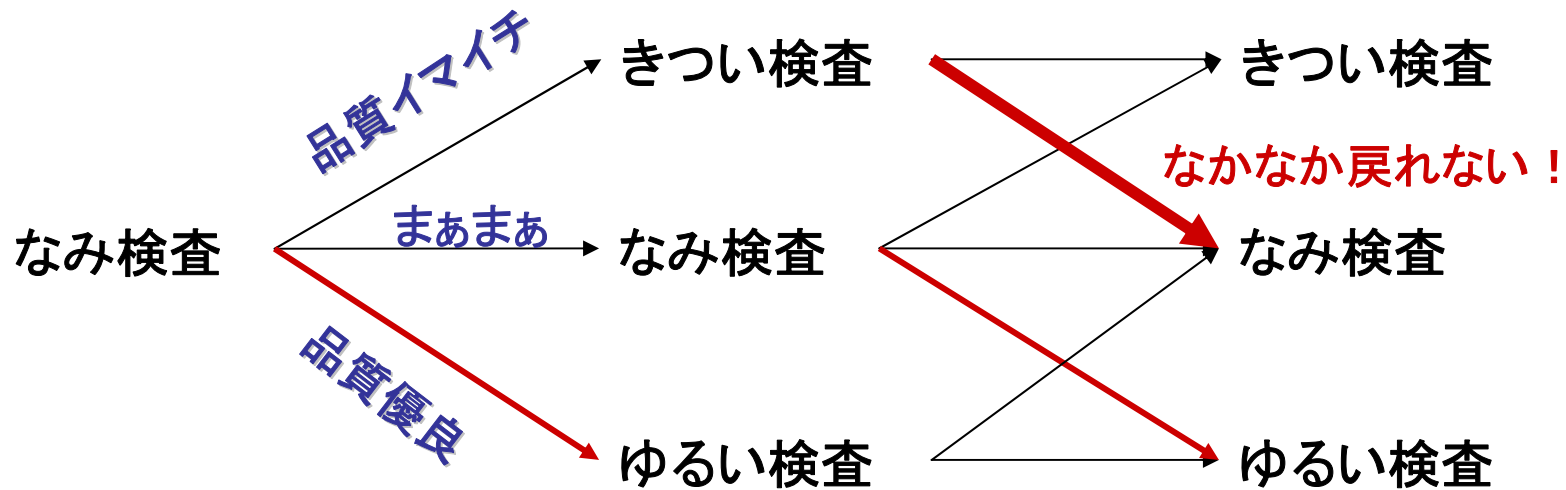
一律にロットサイズNに対して、 $\sqrt{N} + 1$, 10%という基準を使う場合もあるようです。しかし、「統計的な裏付けの無い抜取検査」です。

JIS Z 9015-1の1回抜取方式(なみ検査)

ロットサイズ	サンプル サイズ	AQL			
		1.0%	1.5%	2.5%	4.0%
2 ~ 8	2	1	1	1	1
9 ~ 15	3	1	1	1	1
16 ~ 25	5	1	1	1	1
26 ~ 50	8	1	1	1	2
51 ~ 90	13	1	1	2	2
91 ~ 150	20	1	2	2	3
151 ~ 280	32	2	2	3	4
281 ~ 500	50	2	3	4	6
501 ~ 1200	80	3	4	6	8
1201 ~ 3200	125	4	6	8	11
3201 ~ 10000	200	6	8	11	15
10001 ~ 35000	315	8	11	15	22
35001 ~ 150000	500	11	15	22	22
150001 ~ 500000	800	15	22	22	22
500001 以上	1250	22	22	22	22

これらの数の
エラーが
見つかったら
そのロットは
不適合

JIS Z 9015-1の検査方式



基準に傾斜(キツイ, 並, ユルイ)をつけることによって,
事前に品質の作り込みを行うことを動機づけるという意味がある。

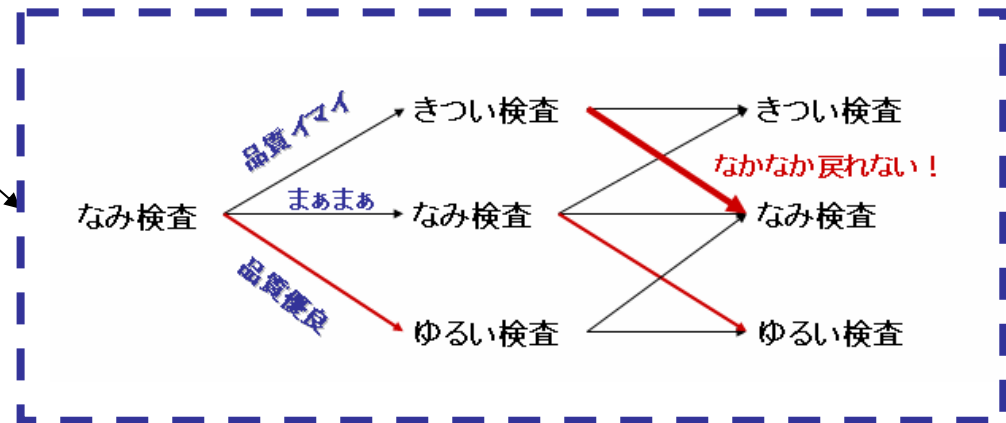
人は後で楽になるなら, なるべくそうなるように努力する。(でしょ?)

●● 他の方式も考えることができる

100%
SDV

100%
SDV

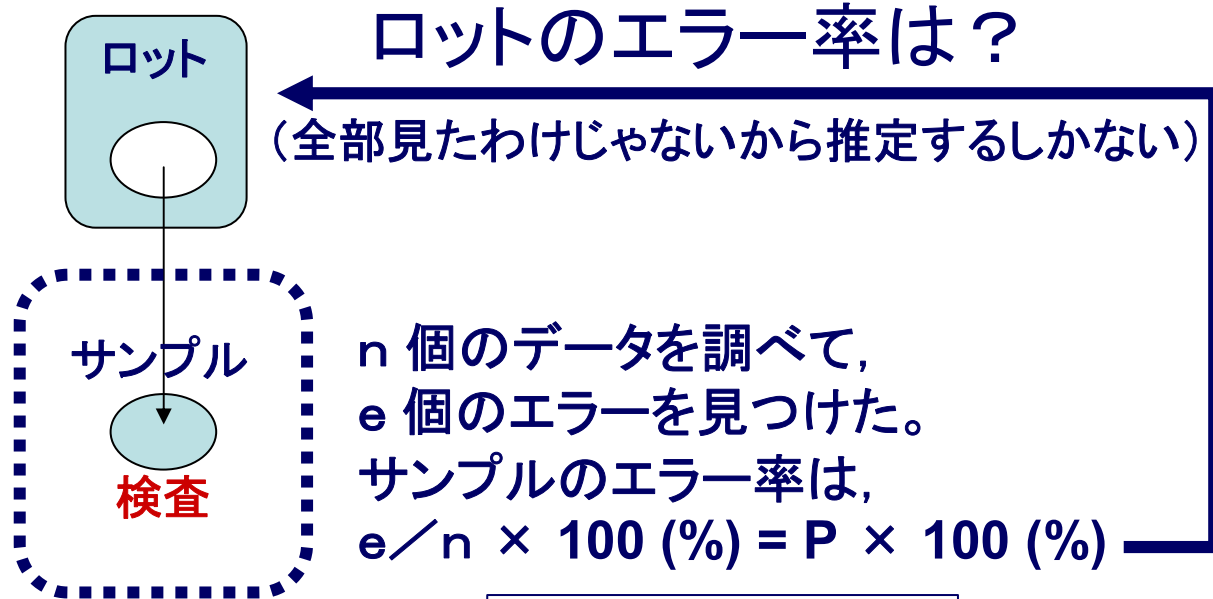
大丈夫なら



サンプリングSDV

施設毎にサンプリングするか否かを決定するとか、
どの施設も最初の症例は100%チェックするとか、
いろいろな方式が考えられる。

●●● ちょっと必要な統計的な考え方



割合 P の95%信頼区間の
大雑把な計算法

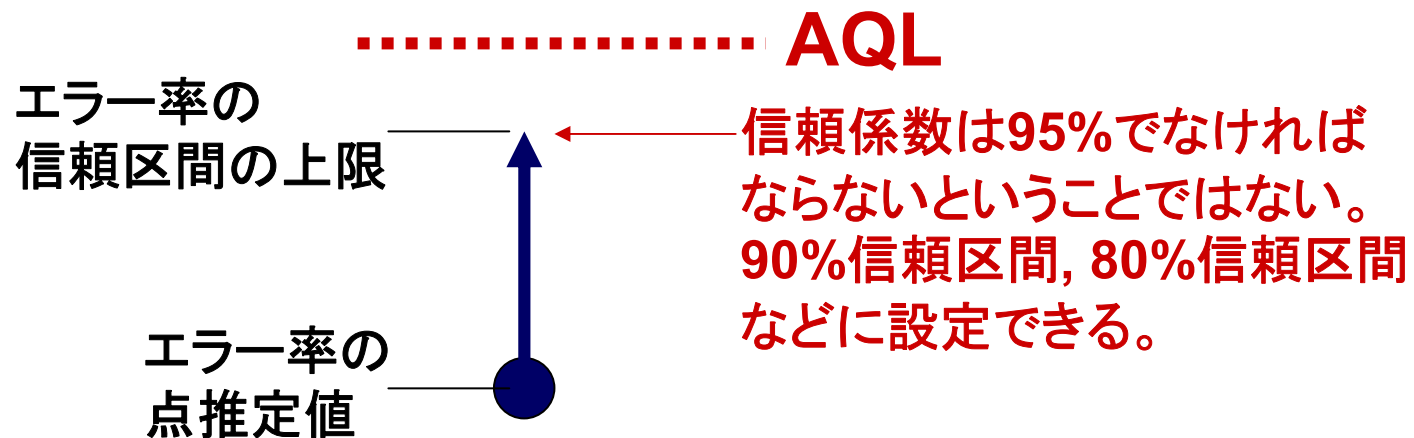
$$2\sqrt{\frac{1}{n} P(1-P)}$$

例えば, 300個のデータ中, 6個のエラーがあったとき,
エラー率の点推定値は $6/300 = 0.02$ で, その95%信頼区間は

$$2\sqrt{\frac{1}{300} \times 0.02 \times 0.98} = 0.016 \quad \text{から, } 0.02 \pm 0.016 \text{ で, 悪くても } 3.6 \%$$

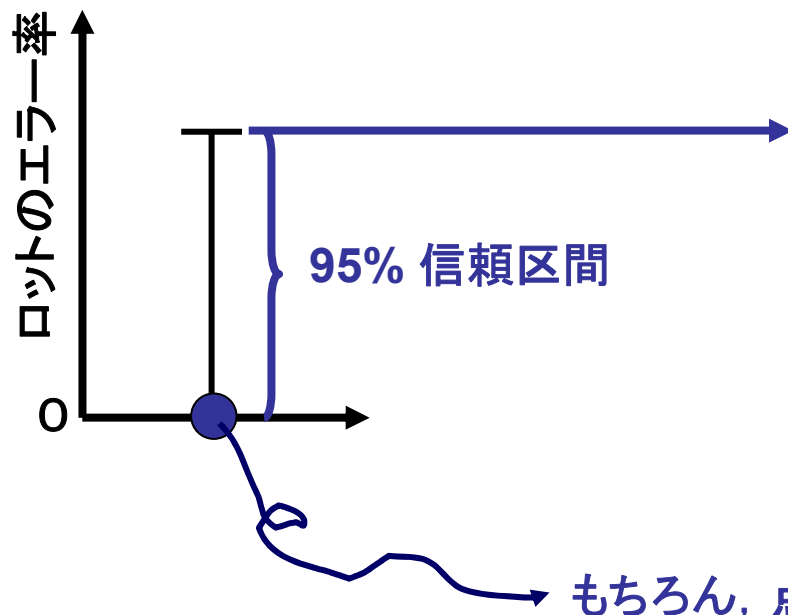
●● AQL とは？

AQL (Acceptable Quality Level; 合格品質限界)
を2% に設定したとする。
これを, 「10000個中200個エラーがあっても許す基準」と
考えるのは, ちょっと違う。
最悪のシナリオを考えても, まずAQLを超えることはないという
基準。



「サンプルにエラーは見つからなかった！」

- N個のデータからなるサンプルをSDVしたらエラーは1つも見つからなかった。
 - サンプルのエラー率は0%→ロット全体のエラー率も0%？
 - ✓ サンプルサイズが小さければエラーは見つかりにくい
 - ✓ サンプルサイズが大きければエラーは見つかりやすい
 - もう一度, N個のデータサンプリングしたら再現できる？



Rule of three (3の法則) !

この上限の値は, $3/N \times 100$ (%)
で近似できる

つまり, 100個データをチェックして,
エラーが見つからなければ,
ロットのエラー率は,
「まあ最悪のシナリオを想定しても3%」
と解釈できる。

もちろん, 点推定値は0%

● ● ●
どのようなサンプリング・プランを考えたとしても、
その動作特性を知っておくことが重要！

- 動作特性＝Operation Characteristics
- その抜取検査が
 - どのくらい不適合を見逃してしまうか
 - どのくらい適合を、誤って不適合とみなしてしまうか

例えば、1, 3, 5, 10, 15, 20, 25, 30例目は100%SDVという状況

6例目までは、「50%のデータを見る」

それ以降は、「20%のデータを見る」に近づいていく

適合・不適合の基準をどのように設定したか？

それはすなわちAQLを何%に設定したことに相当するか？

サンプリングSDVのスキームを考えるときには
必ず統計家も参画すべし

ソビエト的品質管理？

誤りを見つけた者には褒美を，誤りを犯したのものには罰を与える

- 現在でも、品質を作りこむ技術を持たない国あるいは産業分野では、不良品を世の中に出さないことが品質管理の中核となっており、検査重視主義の品質管理がなされている。ISOの第69専門委員会「統計的方法の適用」第4小委員会「統計的工程管理」委員長であったロシアのLapidus (Priority Center所長)は、1994年に「ロシアにおいては、今でも検査重視主義は必要であり、日本のように行かない。品質管理では、一步一步進むしかない」と語った。
- もっとも、Lapidus教授は、これに続けて次のように笑い飛ばした。「ソビエトには、顧客指向の日本的品質管理と並び称すべきソビエト的品質管理システムがあった。第1段階として決して達成不可能な立派な製品規格を発行する。第2段階として、この規格の達成をチェックするために当局は工場を査察する。第3段階として、工場長が当局にとって気に食わないときだけ、当局は工場を規格に違反しているとして摘発する。このシステムを考えついたのはスターリンで、全くスターリンは偉かった。実際、スターリンの管理システムは、アホな為政者でも産業界をコントロール可能なことを実証した。」

統計数理研究所，筑波大学；椿先生から頂いた資料から引用

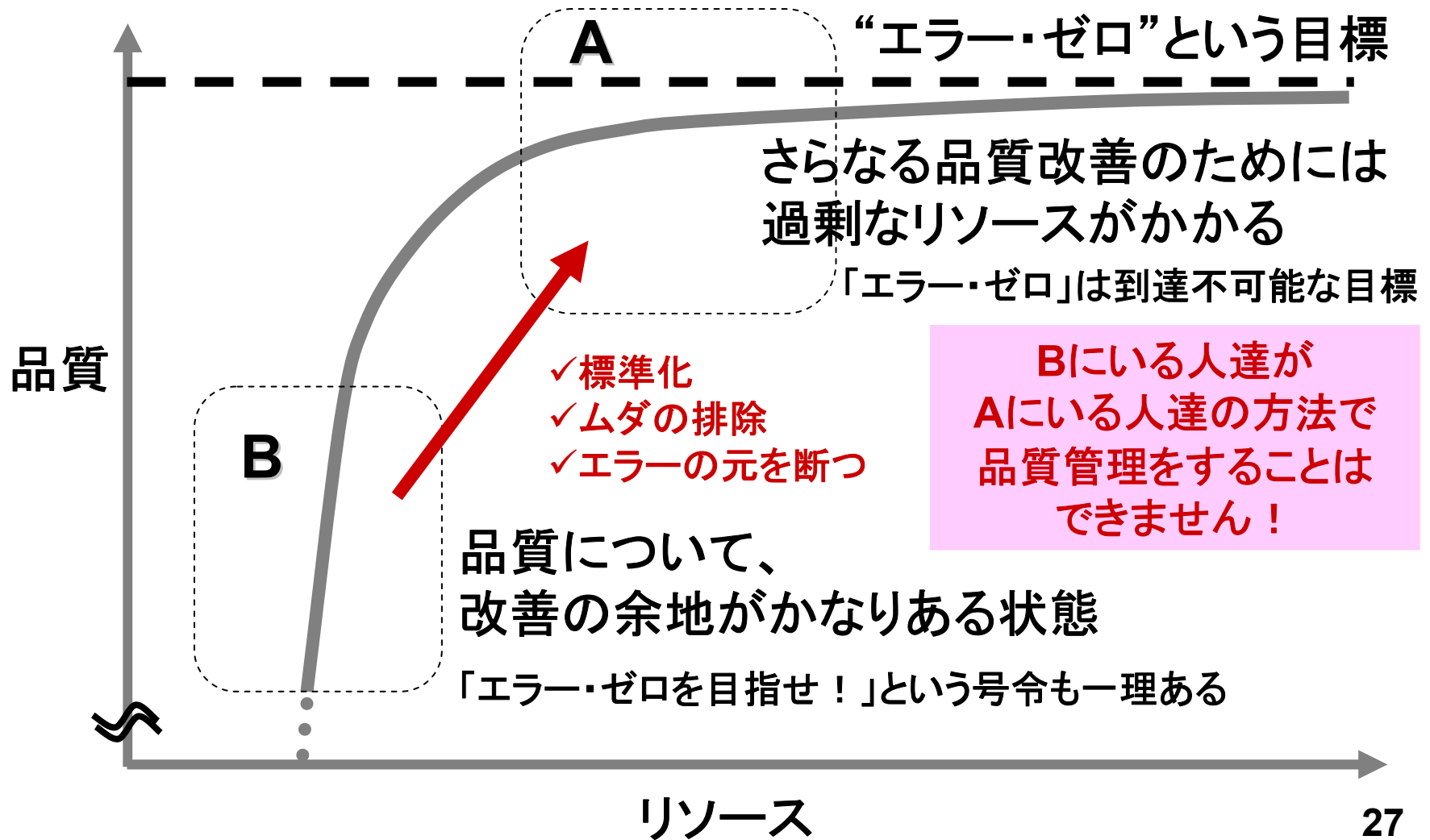
治験データの品質：目指すべき道

- エラーを犯した人を吊るし上げてても根本的な解決にはならない
 - その人だって、不真面目にやってるのではないし、意図的にやってるのでもない。そうさせた原因は何かこそ突き詰めるべき
- 問題はその人にあるのではなく、プロセス、というよりも「プロセスの設計」にある
 - ここで言うプロセスは、患者さんで発生した「データ(情報)」を、収集するための過程すべてを指す
 - スポンサーが要求する「プロセス」もある
 - 医療機関で作っている「プロセス」もある
- プロセスの設計にこそ、スポンサー(CRO)、医療機関が協力して参画すべき
- 一方的に押し付けても、うまく動かない！

スポンサーの視点 vs. PMDAの視点

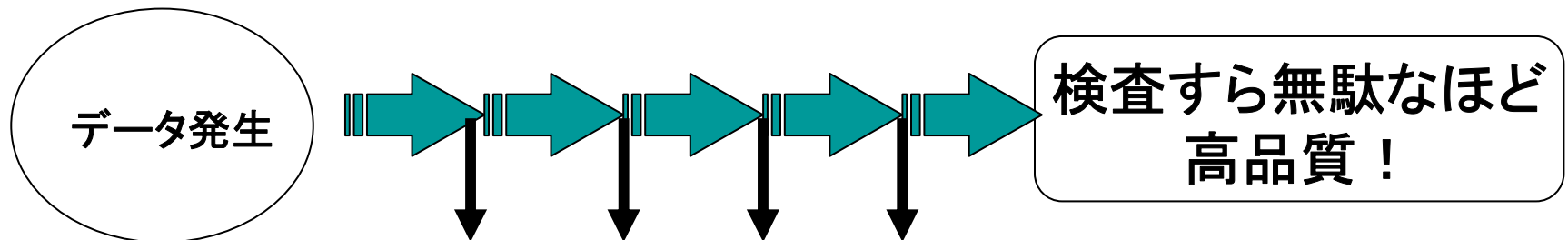
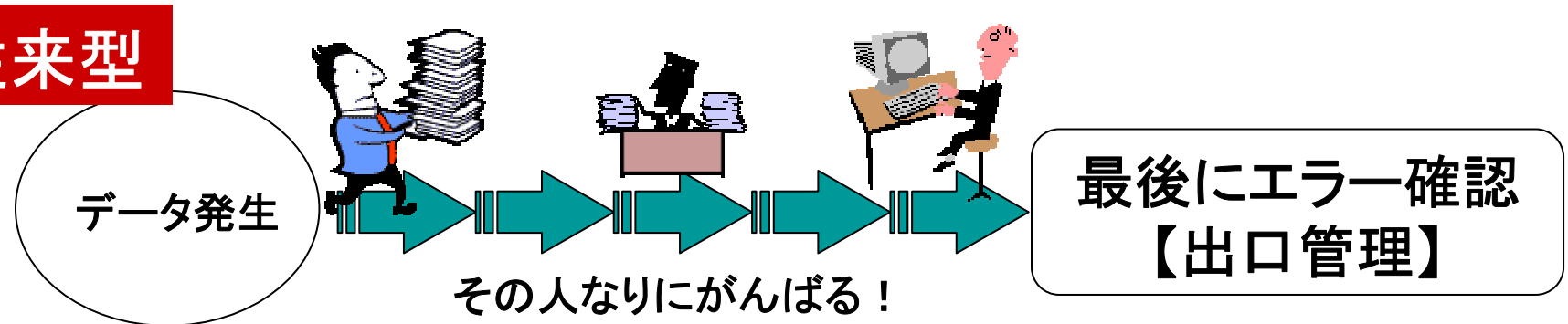
- スポンサーの視点
 - 治験を実施している時に、逐次、品質をモニターしプロセス改善に生かす。
 - でも、顧客(PMDA)がHappyな品質レベルを満たしていることが必要
- PMDAの視点(特に、信頼性保証?)
 - 治験が終わった後で、その治験の結果や規制当局としての意思決定を国民に示したときに、それが、ちゃんとしたデータに基づいていると言えるか(被験者の権利が守られていることは大前提として)

Where are you?



品質確認: 従来型 vs プロセス管理

従来型



プロセスの良さ・悪さを刻々と計測していく

問題があったら、プロセスに悪影響を与える
見逃せない原因 (Assignable Cause) があるはず

プロセス管理

その原因を潰していけば、プロセスはますます安定し
品質を作りこむ力 (Process Capability) は向上する!

シンポ当日あった重要なQ&A

- 【Q】サンプリングでチェックしていないエラーを監査やPMDAから指摘された時の対応は？
 - 【A】そもそも監査やPMDAが個々のエラーを指摘するということがあるとしたら、それこそが「出口管理」です。品質保証のプロセスが上手く機能していたかをチェックするのが、本来のQuality Assuranceのあり方だと思います。致命的なエラーがないかのチェックはしていると思いますが、致命的なエラーがあっては困るデータ・クラスは100%SDVしていればその心配はミニマムなのではないでしょうか？
- 【Q】本当にサンプリングで100%SDVと同等の品質確認ができるのでしょうか？
 - 【A】そのような危惧がある場合には、サンプリングSDVなんて、やっちゃダメです。

Futher Readings

- 「サンプリングSDVへの挑戦」, 小宮山, 多喜田, 臨床医薬 vol.25 no.7, p593-600 (2009)
- 「安全性データの収集と報告に関する提言」, 製薬協 統計・DM部会
- Good Clinical Data Management Practice, Society for Clinical Data Management、統計・DM部会訳(7月刊行予定)
 - 医薬出版センターで購入可能
- JISハンドブック57「品質管理」
- 西堀栄三郎著「品質管理実施法」, 日科技連

西堀栄三郎は、品質管理の専門家というよりは、第1次南極越冬隊長、第1次エベレスト登山隊長、雪山賛歌の作詞者として、日本を代表する冒険家ということで有名。しかし、西堀が行ったのは冒険ではなく、独自の人材操作理論を伴う合理的な課題達成であった。課題達成の天才と位置付けられるべきであろう。主著の一つである「品質管理実施法」(日科技連)は、シューハートの工程管理、工程解析を実際の問題解決に活かす知恵が詰まっている。