



国交省・下水サーベイランスデータ標準化の試み

遠藤礼子, PhD

京都大工学研究科附属流域圏総合環境質研究センター
招聘研究員

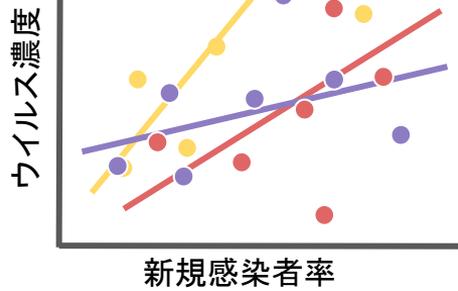
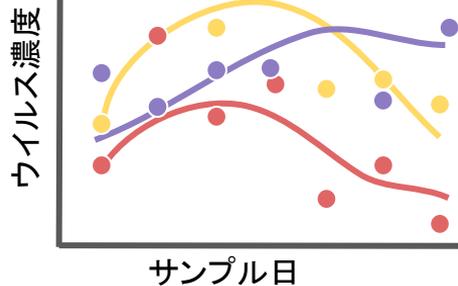


分析1次データ

- A処理場;a分析法
- B処理場;b分析法
- C処理場;b分析法

分析手法、希釈率
等が異なると分析1
次データの量的な
相互比較ができて
ない

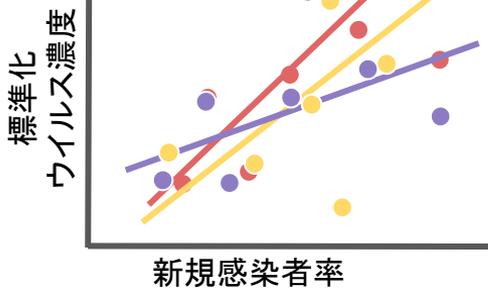
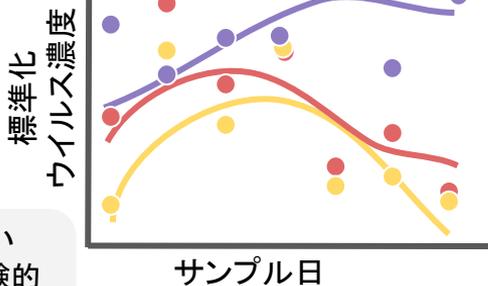
規格化されない
データから経験的
・統計的法則を導
くのは困難



規格化 (Regularization)

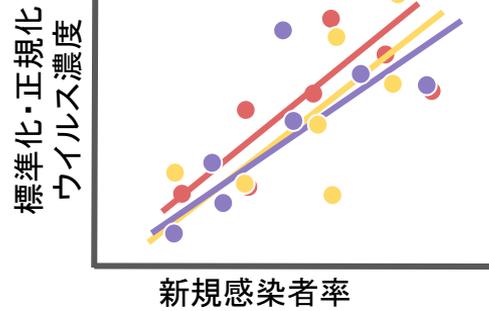
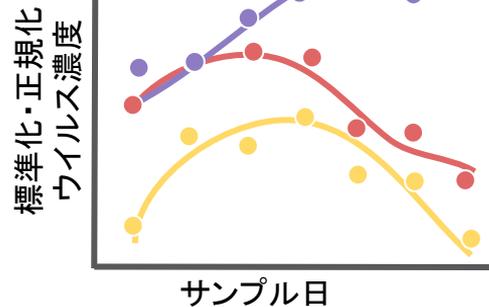
標準化 (Standardization)

手法による回収率の補正
(手法間のバイアスの補正)

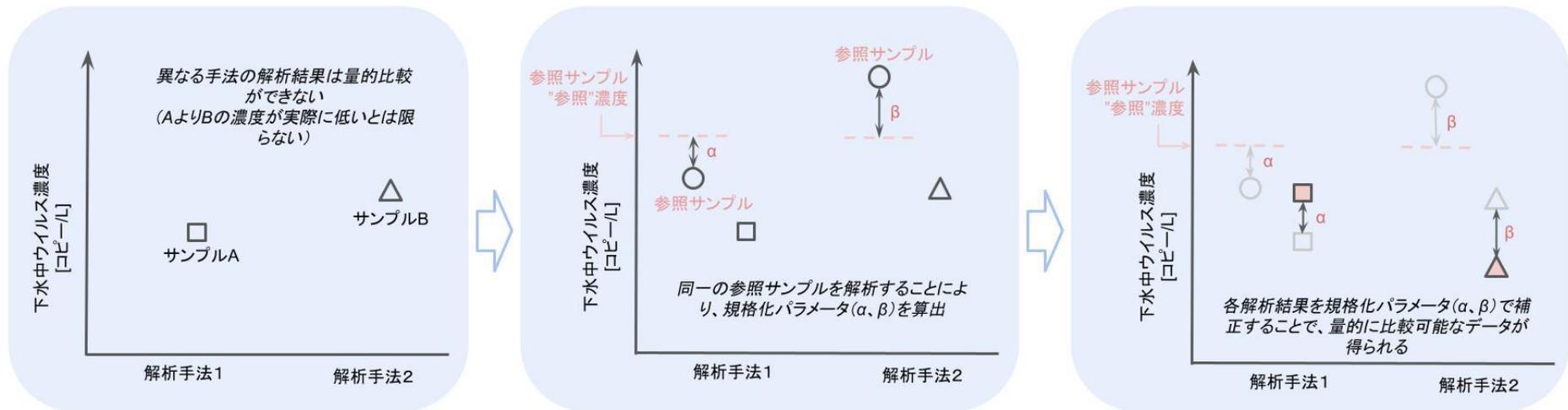


正規化 (Normalization)

合流、工場排水、人口密度等による希釈率
等の補正 (サンプル間のバラツキの補正)



日本版下水サーベイランスデータ標準化手法の提案





標準化試験概要：メタデータ

標準化試験の対象分析手法・分析機関を右に、試験日程を下に示す。

令和4年12月12-13日：

参照サンプルの準備（下水サンプル取得、参照サンプルの作成、サンプルキットの送付）

令和4年12月14-23日：

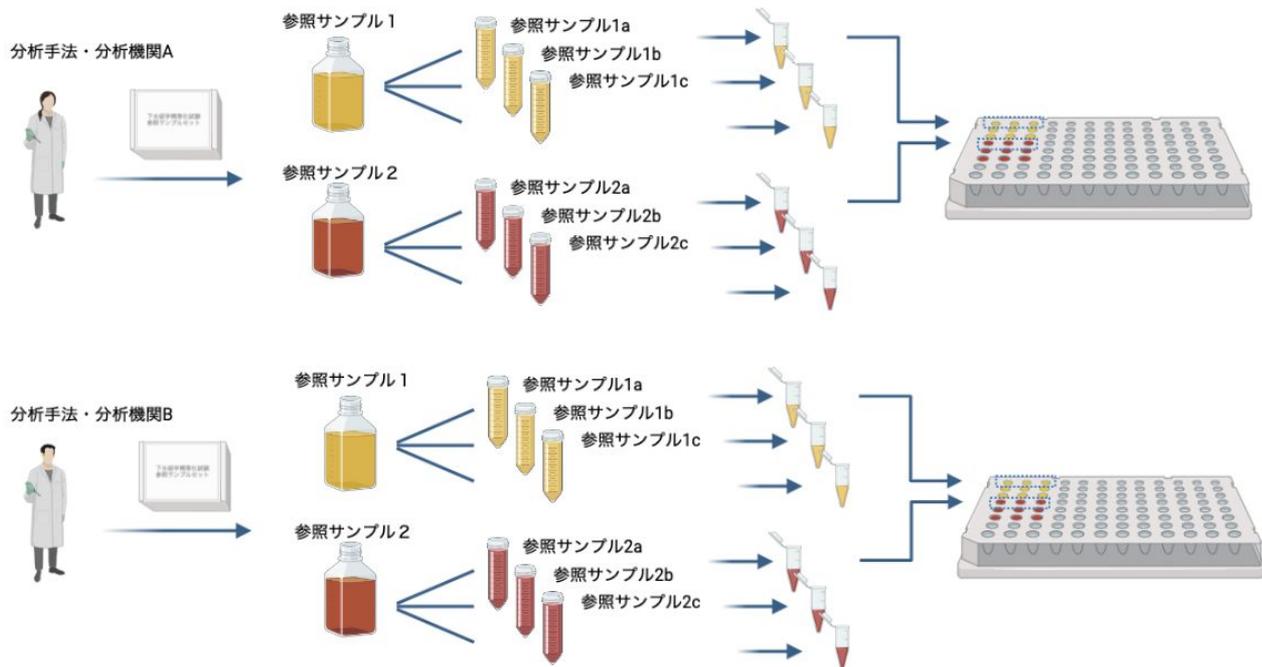
対象となる分析機関にて参照サンプルの分析を実施

令和4年12月26-令和5年1月6日：

標準化試験の結果の分析とまとめ

サンプル対象	分析手法	分析機関	分析対象
下水上清	PEG沈殿法	島津テクニサーチ	N1、N2、PMMoV
下水沈殿物	沈殿物抽出法	島津テクニサーチ	N1、N2、PMMoV
下水沈殿物	沈殿物抽出法	北里環境科学センター	N1/N2同時検出、PMMoV
下水沈殿物	EPISENS・S法	塩野義	N1、PMMoV
下水沈殿物	EPISENS・S法	北海道大学	N1、PMMoV
下水沈殿物	EPISENS・M法	北海道大学	N1、PMMoV

標準化試験概要：試験デザイン



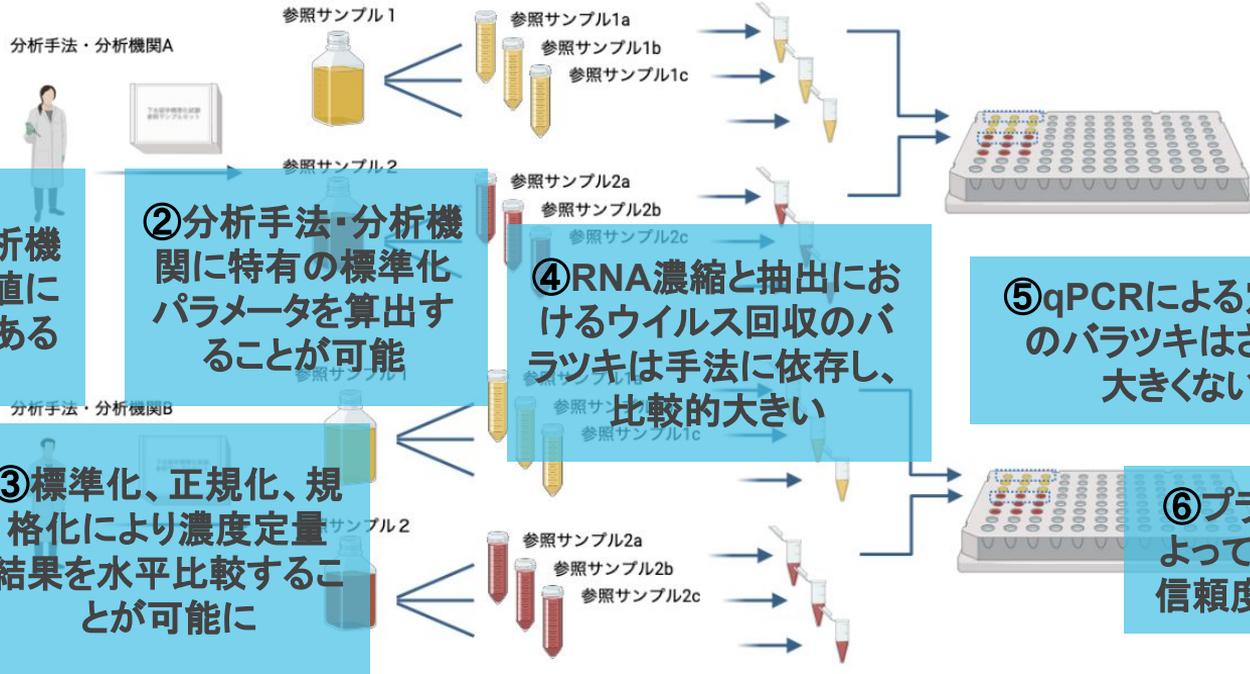
分析手法・分析機関それぞれに標準化試験サンプルキットを送付。

参照サンプルは2種類あり、それぞれを3複製の計6つのサンプル (2つの *biological replicates* x 3つの *technical replicates*) がある。

6つのサンプルはそれぞれRNA濃縮・抽出を行う。

抽出されたRNAはそれぞれ3ウェル分ウイルス濃度を測定。
SARS-CoV-2 N1 及びN2、PMMoVをターゲットに定量化し、Ct値及び濃度を算出する。

標準化試験概要：結果のまとめ



分析手法・分析機関それぞれに標準化試験サンプルキットを送付。

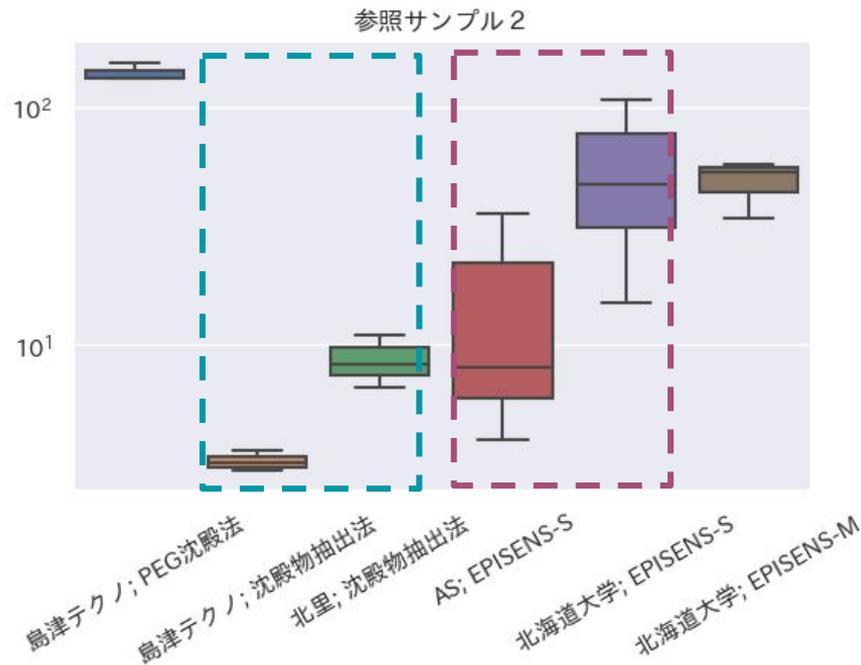
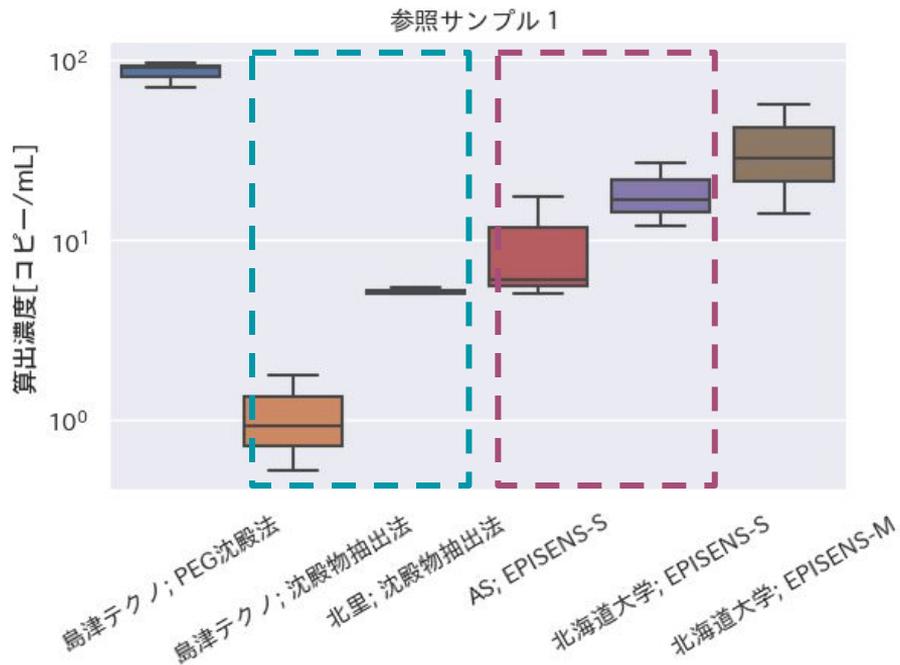
参照サンプルは2種類あり、それぞれ3複製の計6つのサンプル (2つの *biological replicates* x 3つの *technical replicates*) がある。

6つのサンプルはそれぞれRNA濃縮・抽出を行う。

抽出されたRNAはそれぞれ3ウェル分ウイルス濃度を測定。
SARS-CoV-2 N1 及びN2、PMMoVをターゲットに定量化し、Ct値及び濃度を算出する。

①分析手法・分析機関によって定量値にかなりの違いがある

同じ手法を使っていたとしても、分析機関によって違いがある

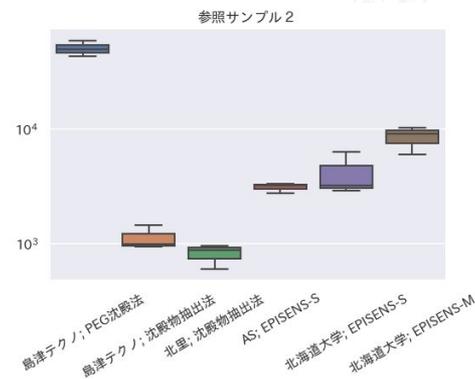
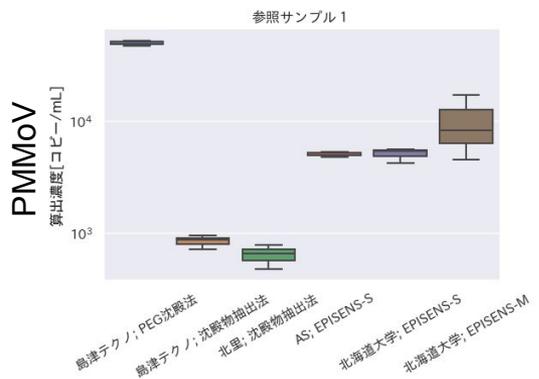
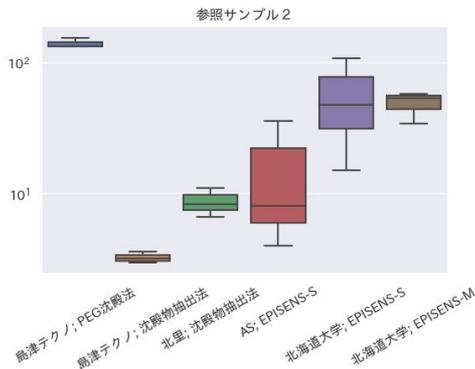
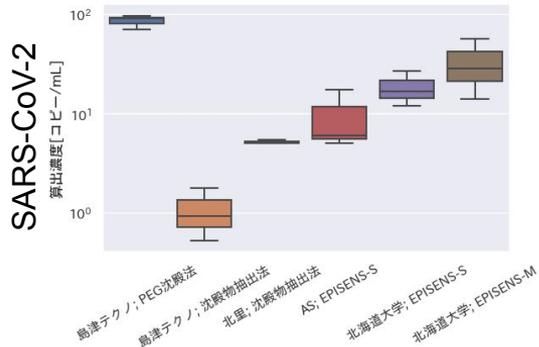


注: 島津テクノ両メソッドの結果はN2プライマーに基づくもののみを使用

標準化試験結果の妥当性：論文との比較

標準化試験の結果

その他の独立した論文より



PEG沈殿法に比べたEPISENS法の濃度値比較
([Ando et al. \(2022\) Sci.Total Environ.](#))

- SARS-CoV-2: オーダーが1~2低い ✓
- PMMoV: さほど変わらない ✓

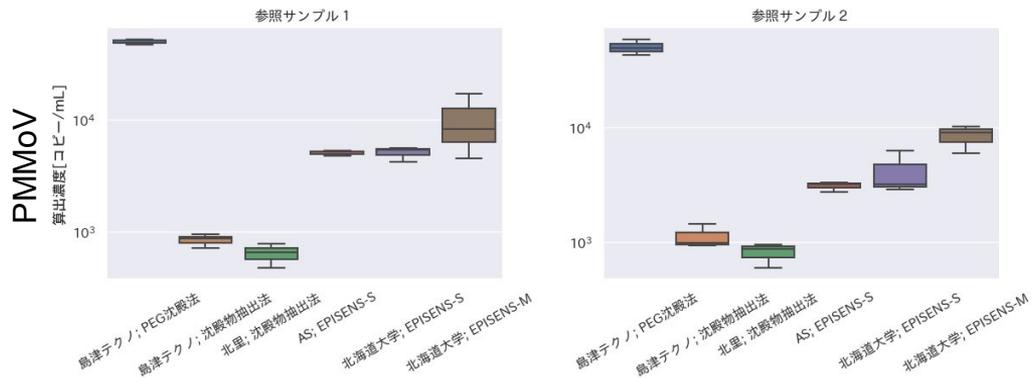
PEG沈殿法に比べた沈殿物抽出法の濃度値比較
([Kitamura et al. \(2021\) Sci.Total Environ.](#))

- SARS-CoV-2: 十分なデータなし □
- PMMoV: オーダーが1~2低い ✓

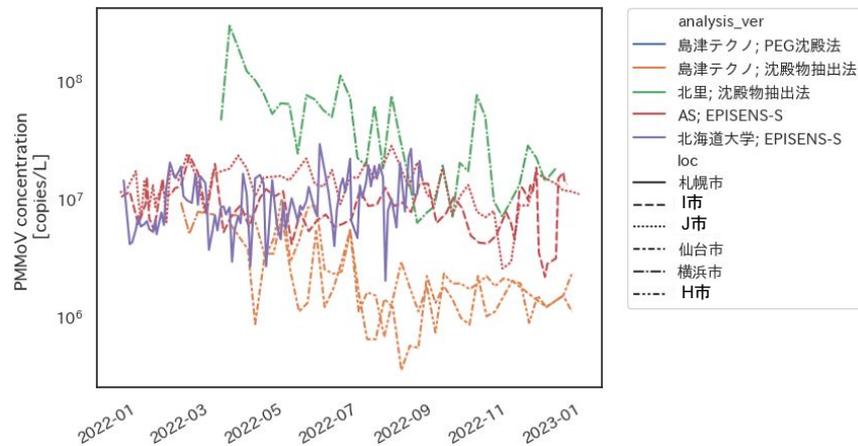
→概して妥当な結果である

標準化試験結果の妥当性: PMMoVフィールドデータとの比較

標準化試験の結果



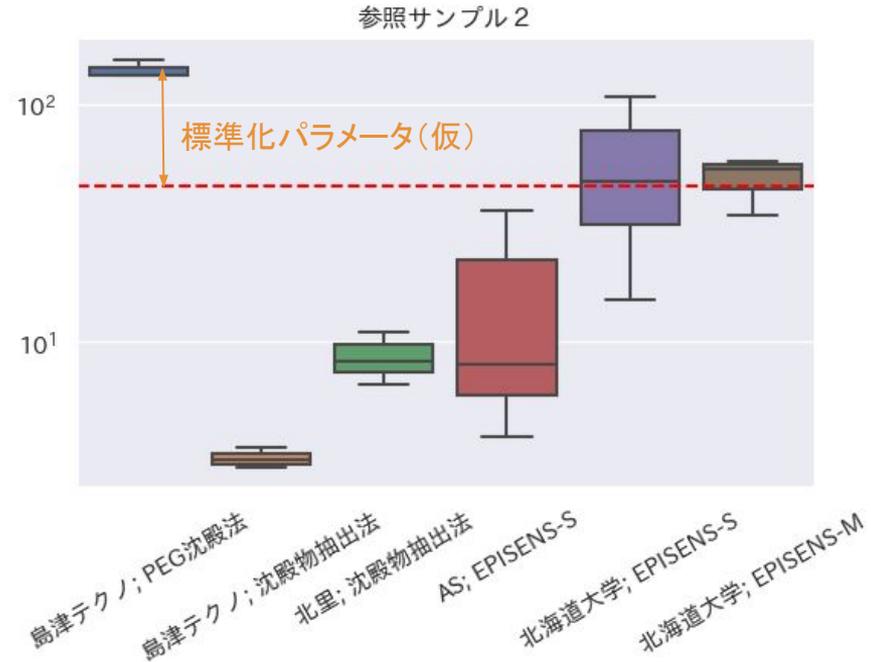
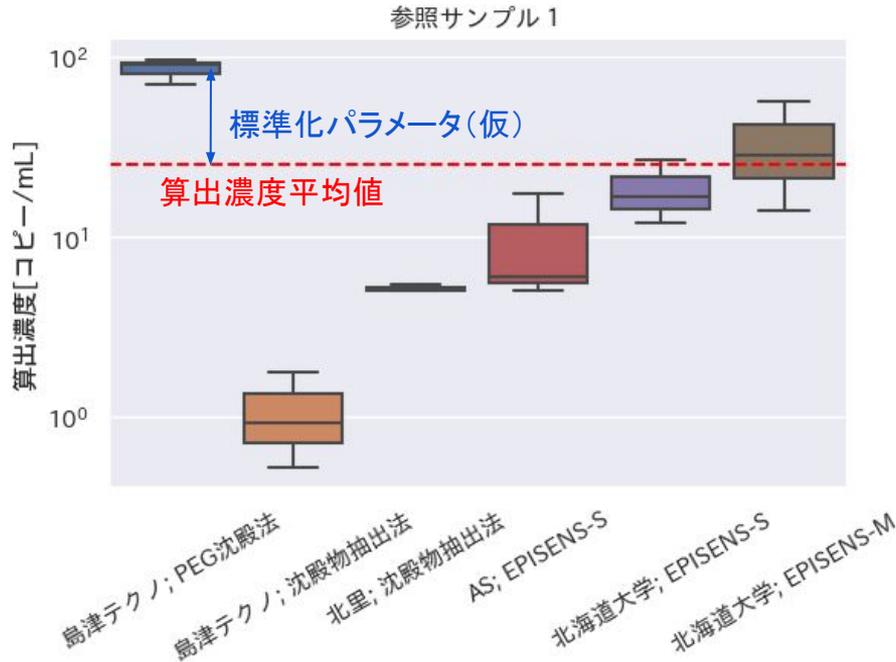
フィールドデータ



仮に希釈率が一定だとすると、PMMoV濃度は流行動向に関係なく、標準化試験の結果に従うと考えられるが、

- ・北里; 沈殿物抽出法の濃度が想定以上に高い
- ・その他の手法に基づくデータの挙動はおおむね想定通り。

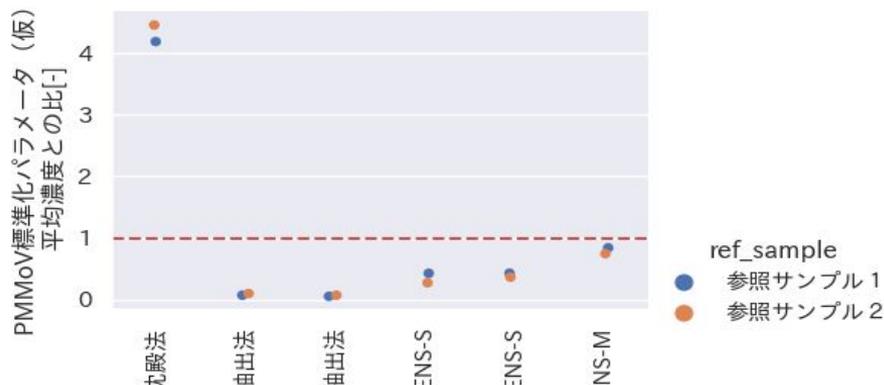
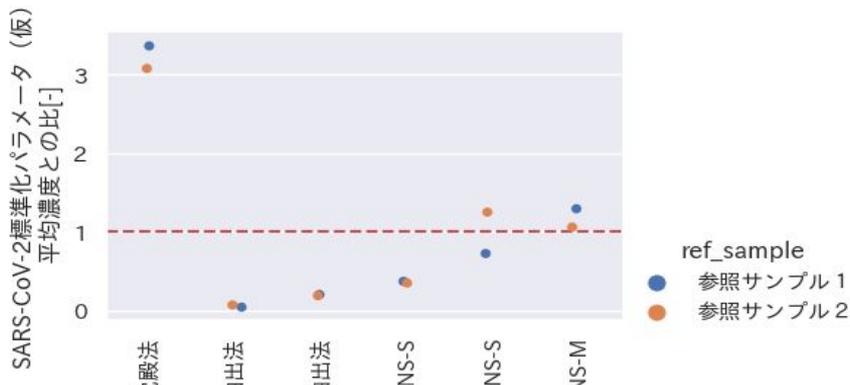
定量値のバイアスは分析手法・分析機関に特有で、サンプルに依存しないか？→もしそうなら標準化パラメータでデータの標準化が可能に



注: 島津テクノ両メソッドの結果はN2プライマーに基づくもののみを使用

②分析手法・分析機関に特有の標準化パラメータを算出することが可能

「試験全体での各参照サンプル濃度平均に対するおのこの分析手法・分析機関で算出した濃度の比」をそれぞれの分析手法・分析機関、そして参照サンプルごとに算出した。この値は、分析手法・分析機関に特有で、サンプルへの依存度が少ないことから、おのこの分析手法・分析機関がウイルス濃度をどれだけ過大 / 過小評価しやすいかは、下水サンプルによらず、ある程度一定であることが示された。このことは、各分析手法・分析機関に対して標準化パラメータ(定数)を定めればそれぞれの分析機関で算出される濃度の定量値を統一的に比較できるようになることを示す。「試験全体での濃度平均に対するおのこの分析手法・分析機関で算出した濃度の比」などが標準化パラメータとして使用されうることを示す。

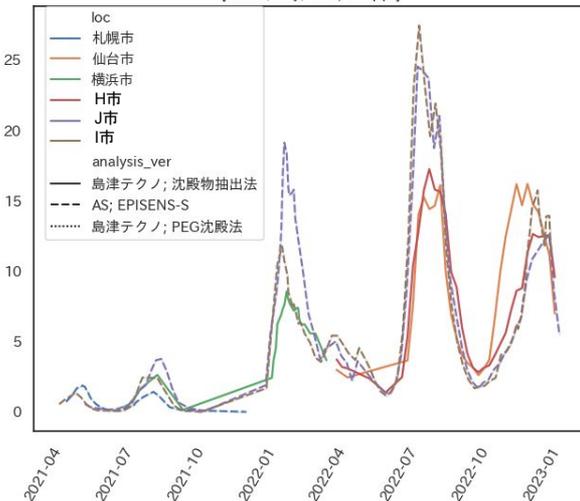


→ 標準化パラメータ(仮)の平均を標準化パラメータとして、実際に分析されたフィールドデータで標準化を実行してみる

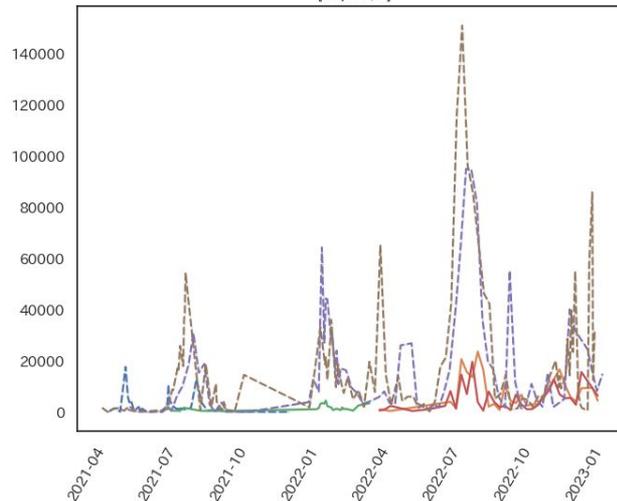


フィールドデータとの比較（2手法、4-6都市）

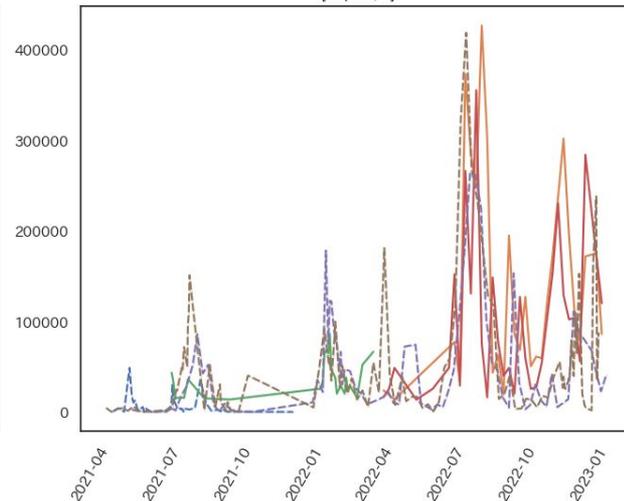
10万人あたり新規感染者数の1週間平均
[cases/day/100,000pp]



SARS-CoV-2濃度 (補正なし)
[copies/L]

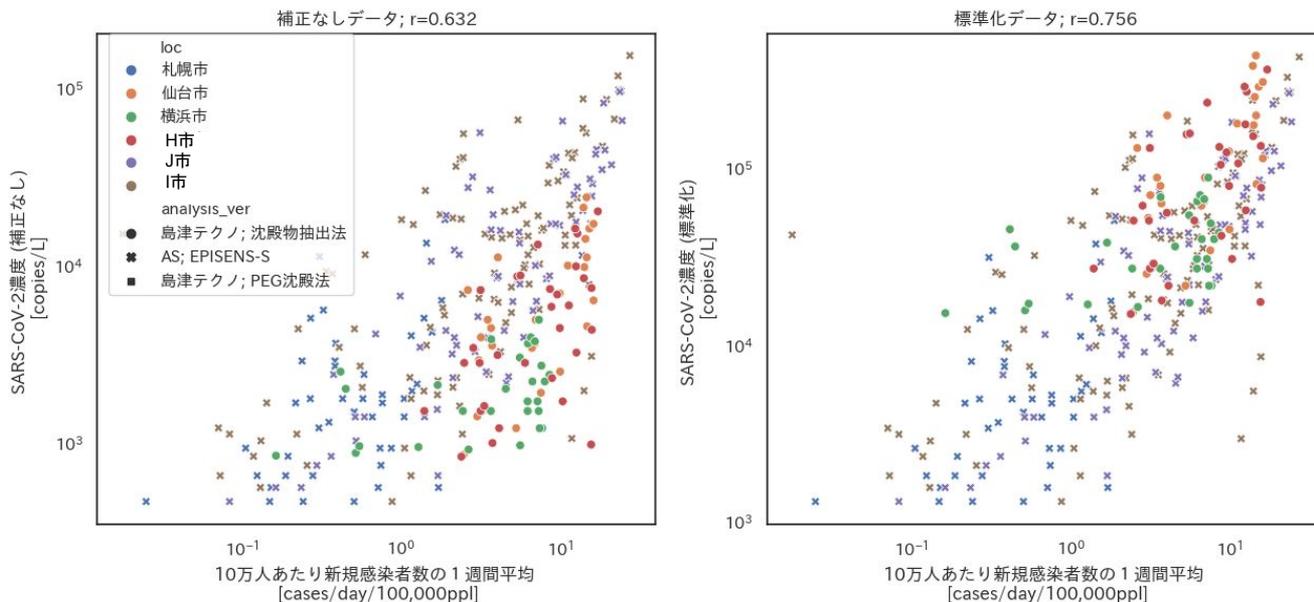


SARS-CoV-2濃度 (標準化)
[copies/L]



島津テクノ; 沈殿物抽出法とAS ;EPISENSのみデータの標準化を適用。
データ標準化を行うことにより、下水中ウイルス濃度の都市間トレンドは、新規感染者率の都市間
トレンドに近づいた。

フィールドデータとの比較（2手法、4-6都市）



島津テクノ;沈殿物抽出法とAS ;EPISENSのみデータの標準化を適用。
データ標準化を行うことにより、下水中ウイルス濃度の都市間トレンドは、新規感染者率の都市間トレンドに近づいた。

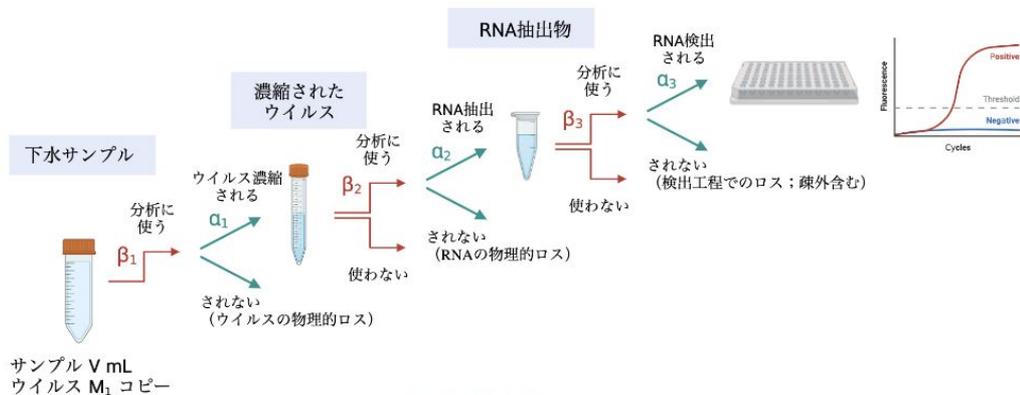
注意点：標準化パラメータ及び標準化試験での濃度定量値は手法の良し悪しを判断するものではない

手法の良し悪しを図るためには以下のような複数の観点が必要。

- 分析感度 (sensitivity) が高く、低濃度でも検知が可能
 - 高い回収率・低い阻害要因 (α)
 - 高い分析液量率 (β)
 - 多くのサンプルボリューム (V)
- 分析精度 (precision) が高く、ブレの少ない測定が可能
 - (α, β) では必ずしも判断できない
- 分析にかかる時間が短い
- 安価

etc.

=> 標準化パラメータはプロセス効率 (α) を近似するものだが、それだけで手法の良し悪しを判断することはできない。また、標準化試験により算出された濃度の大小でも手法の良し悪しを図ることはできない。



通常報告されるサンプル中のウイルス数 (M_2) ・濃度 (C_2)
(プロセス効率を無視)

$$M_2 = M_1 \times \alpha = M_3 / \beta$$

$$C_2 = M_2 / V$$

回収率
(= 1 - ロス率)

$$\alpha = \frac{\alpha_1 \times \alpha_2 \times \alpha_3}{\alpha_3}$$

実際のサンプル中のウイルス数 (M_1) ・濃度 (C_1)

$$M_1 = M_2 / \alpha = M_3 / (\alpha \times \beta)$$

$$C_1 = M_1 / V$$

分析液量率
(= PCR反応に供する下水相当量 / V)

$$\beta = \beta_1 \times \beta_2 \times \beta_3$$

検出されたウェル中のウイルス数 (M_3)

$$M_3 = M_2 \times \beta$$

$$= M_1 \times (\alpha \times \beta)$$

$$= C_1 \times (\alpha \times \beta \times V)$$



以下参考スライド

標準化試験概要：参照サンプル





データ分析と統計処理

各分析手法・分析機関が各ウェルごとに算出した下水中ウイルス濃度 [コピー/mL]: $c(i,j,k,l)$

i : 分析手法・分析機関の組み合わせ、 $i = 1, 2, \dots, 6$

j : 参照サンプルタイプ、 $j = 1, 2$

k : 参照サンプルの technical replicates、 $k = 1, 2, 3$

l : qPCRウェルでの繰り返し分析、 $l = 1, 2, 3$

各分析手法・分析機関が各検体 (参照サンプルの technical replicates) ごとに算出した下水中ウイルス濃度 [コピー/mL]: $c(i,j,k)$

$$c(i, j, k) = \frac{1}{3} \sum_{l=1}^3 c(i, j, k, l)$$

各分析手法・分析機関が参照サンプルごとに算出した下水中ウイルス濃度 [コピー/mL]: $c(i,j,k)$

$$c(i, j) = \frac{1}{3} \sum_{k=1}^3 c(i, j, k)$$

参照サンプルごとの下水中ウイルス濃度、全ての分析手法・分析機関による平均 [コピー/mL]: $c(j)$

$$c(j) = \frac{1}{6} \sum_{i=1}^6 c(i, j)$$

データ分析と統計処理

分析手法・分析機関ごとの標準化パラメータ [-];

$$\alpha(i) = \frac{1}{2} \sum_{j=1}^2 \frac{c(i,j)}{c(j)}$$

標準化した下水中ウイルス濃度 [コピー/mL]: $c_{\text{standardized}}(i,j,k)$

$$c_{\text{standardized}}(i,j,k) = \frac{c(i,j,k)}{\alpha(i)}$$

正規化した下水中ウイルス濃度 [コピー/mL]: $c_{\text{normalized}}(i,j,k,l)$

$$c_{\text{normalized}}(i,j,k) = c(i,j,k) / \frac{c^{\text{PMMoV}}(i,j,k)}{c^{\text{PMMoV}}(i)}$$

$$\overline{c^{\text{PMMoV}}}(i) = \frac{1}{6} \sum_{j=1}^2 \sum_{k=1}^3 c(i,j,k)$$

規格化した下水中ウイルス濃度 [コピー/mL]: $c_{\text{regularized}}(i,j,k,l)$

$$c_{\text{regularized}}(i,j,k) = \frac{c(i,j,k)}{\alpha(i)} / \frac{c^{\text{PMMoV}}(i,j,k)}{c^{\text{PMMoV}}(i)}$$





データ分析と統計処理

バラツキを評価する分析手法・分析機関ごとの濃度が大きく異なるため、バラツキは標準偏差ではなく、標準偏差を平均で割った変動係数 (Coefficient of Variation または Normalized Standard Deviation) で評価した。

分析手法・分析機関ごとのRNA濃縮・抽出時のウイルス回収のバラツキ
[-]: $CV_k(i)$

$$CV_k(i) = \frac{1}{2} \sum_{j=1}^2 \frac{\sigma_k(i,j)}{\mu_k(i,j)}$$

$$\sigma_k(i,j) = \sqrt{\frac{1}{3} \sum_{k=1}^3 (c(i,j,k) - c(i,j))^2}$$

$$\mu_l(i,j) = c(i,j)$$

分析手法・分析機関ごとのqPCR定量化のバラツキ[-]: $CV_l(i)$

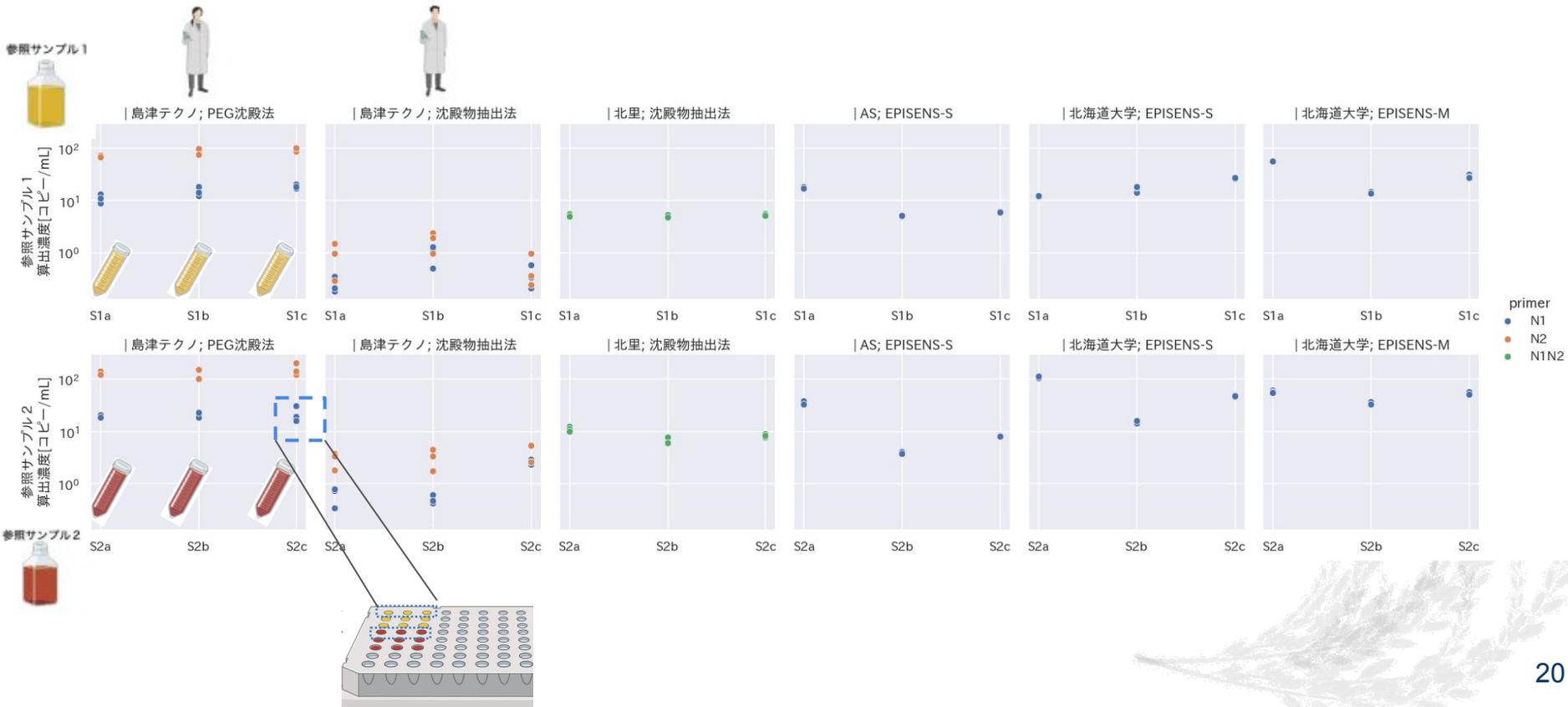
$$CV_l(i) = \frac{1}{6} \sum_{j=1}^2 \sum_{k=1}^3 \frac{\sigma_l(i,j,k)}{\mu_l(i,j,k)}$$

$$\sigma_l(i,j,k) = \sqrt{\frac{1}{3} \sum_{l=1}^3 (c(i,j,k,l) - c(i,j,k))^2}$$

$$\mu_l(i,j,k) = c(i,j,k)$$



試験結果(生データ)



④RNA濃縮と抽出におけるウイルス回収のバラツキは手法に依存し、比較的大きい

6サンプルそれぞれにつき RNA抽出液 3ウェル分の定量結果の平均を取り、その結果の変動係数を参照サンプル1、参照サンプル2ごとに算出。その2つの標準偏差の平均をとった。

分析手法	分析機関	プライマー	変動係数	
			Ct値	濃度
PEG沈殿法	島津テクノロジー	N1	0.00701	0.160
		N2	0.00527	0.119
		PMMoV	0.00694	0.107
沈殿物抽出法	島津テクノロジー	N1	0.0323	0.847
		N2	0.0153	0.343
		PMMoV	0.0113	0.196
沈殿物抽出法	北里環境科学センター	N1N2	0.00672	0.150
		PMMoV	0.0142	0.234
EPISENS-S	塩野義	N1	0.0523	0.905
		PMMoV	0.00418	0.0756
EPISENS-S	北海道大学	N1	0.0408	0.627
		PMMoV	0.0161	0.300
EPISENS-M	北海道大学	N1	0.0298	0.455
		PMMoV	0.0290	0.457

⑤qPCRによる定量化のバラツキはさほど大きくない

RNA抽出液を3ウェル分 qPCRで分析した際の定量値のバラツキを測定した。6サンプルごとに3ウェル分の測定値の変動係数を算出し、その6サンプル分の平均をとった。

分析手法	分析機関	プライマー	変動係数	
			Ct値	濃度
PEG沈殿法	島津テクノロジー	N1	0.00683	0.168
		N2	0.00671	0.143
		PMMoV	0.00607	0.0930
沈殿物抽出法	島津テクノロジー	N1	0.0143	0.340
		N2	0.0229	0.511
		PMMoV	0.00464	0.0796
沈殿物抽出法	北里環境科学センター	N1N2	0.00320	0.0728
		PMMoV	0.00250	0.0432
EPISENS-S	塩野義	N1	0.00215	0.0383
		PMMoV	0.00274	0.0502
EPISENS-S	北海道大学	N1	0.00309	0.0531
		PMMoV	0.00419	0.0747
EPISENS-M	北海道大学	N1	0.00311	0.0515
		PMMoV	0.00377	0.0620

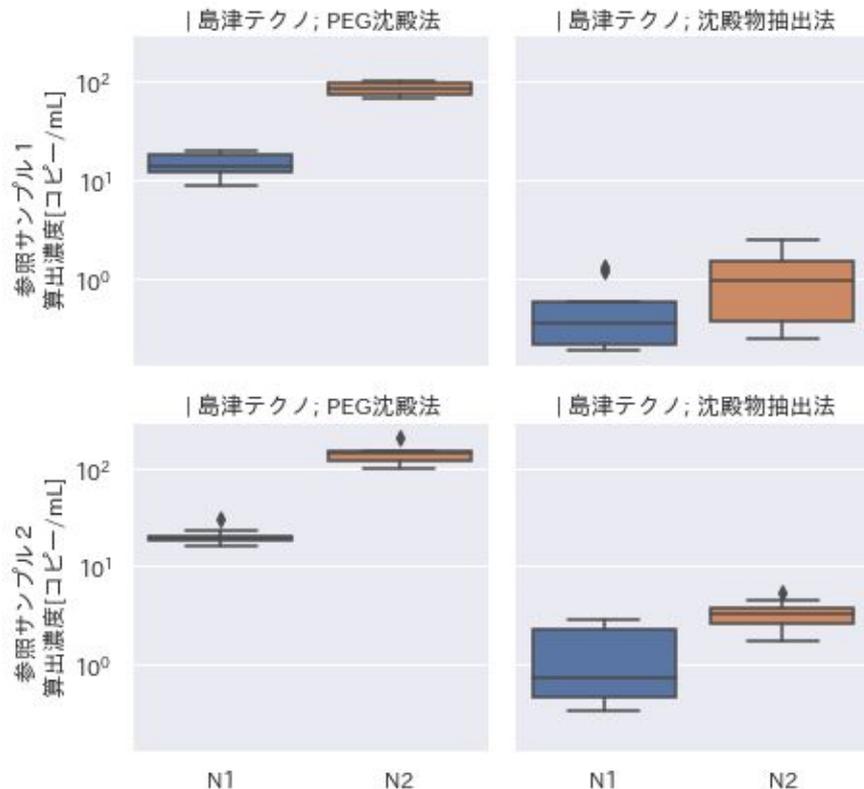


⑥プライマーによって定量化の信頼度が異なる

SARS-CoV-2の定量化をN1、N2を両方用いてそれぞれ定量化した2つの手法(島津テクニサーチによる PEG沈殿法と同分析会社による沈殿物抽出法)に対し、プライマーの違いにより SARS-CoV-2の定量化に与える影響を検討した。

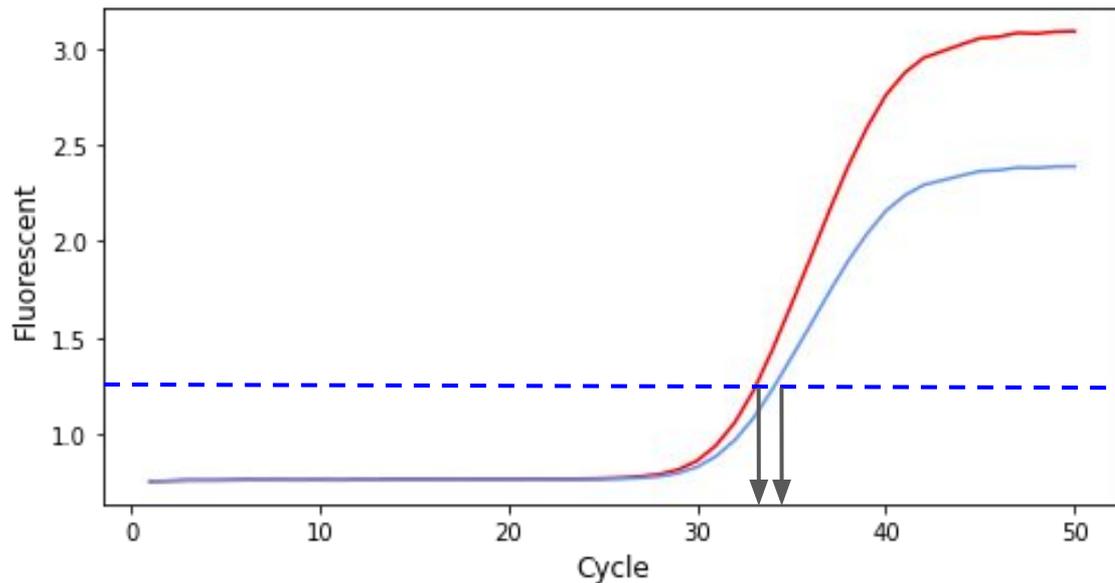
2種類の参照サンプルを用いて比較をしたところ、N2プライマーを用いたSARS-CoV-2定量値の結果に対して、N1プライマーを用いた定量値の結果は、PEG沈殿法では約15%ほど、沈殿物抽出法では約40%ほどしか検出されなかった。

オミクロン株にN1プローブ位置への変異が現れて以来、N1プライマーを用いた際のPCR増幅効率の低下が確認されている。本標準化試験においてもPEG沈殿法、沈殿物抽出法共にN1プライマーを使った際のPCR増幅効率の低下が原因とみられる濃度の過小評価がみられた。なお、N1プライマーを用いた際の濃度の過小評価の程度は分析手法によること、また同じ分析手法でもPCRマスターミックスの種類などによっても異なることが本標準化試験以外の研究(論文未発表)において知られている。



(補足) N1エリアにおける変異によりqPCR定量化に与える影響

サンプル解析時にqPCRから吐き出される増幅曲線

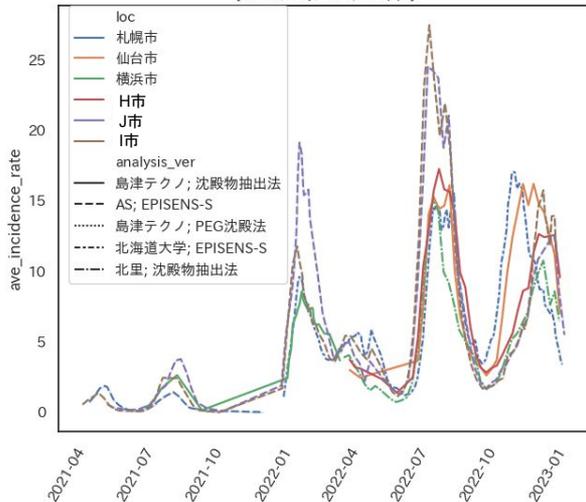


- 1、プローブに出現した変異によって増極効率が落ちる(赤 → 青)。
- 2、CP法を用いて定量化した場合は Ct値が大きくなる。
なお、SDM法やCy0法はこの影響を受けにくい。

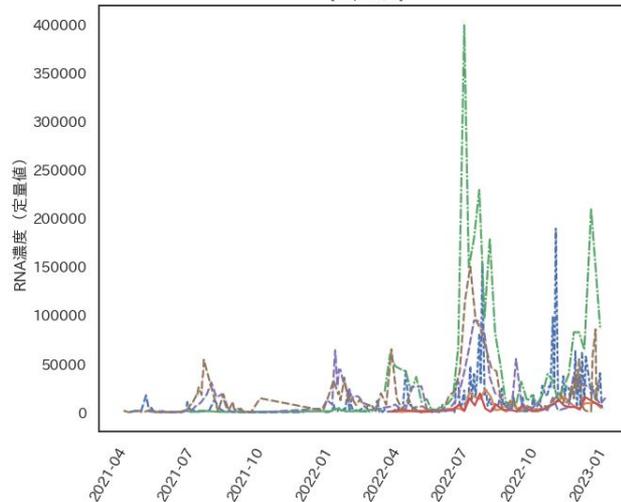


フィールドデータとの比較：罹患者率との比較（全手法）

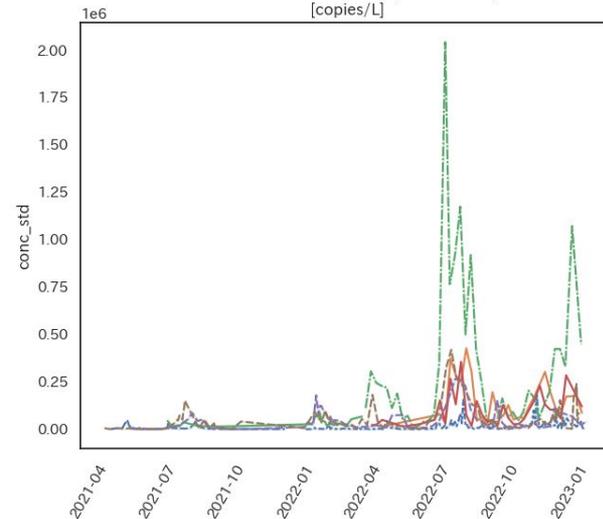
Average incidence rate
[cases/day/100,000pp]



SARS-CoV-2 concentration (raw)
[copies/L]

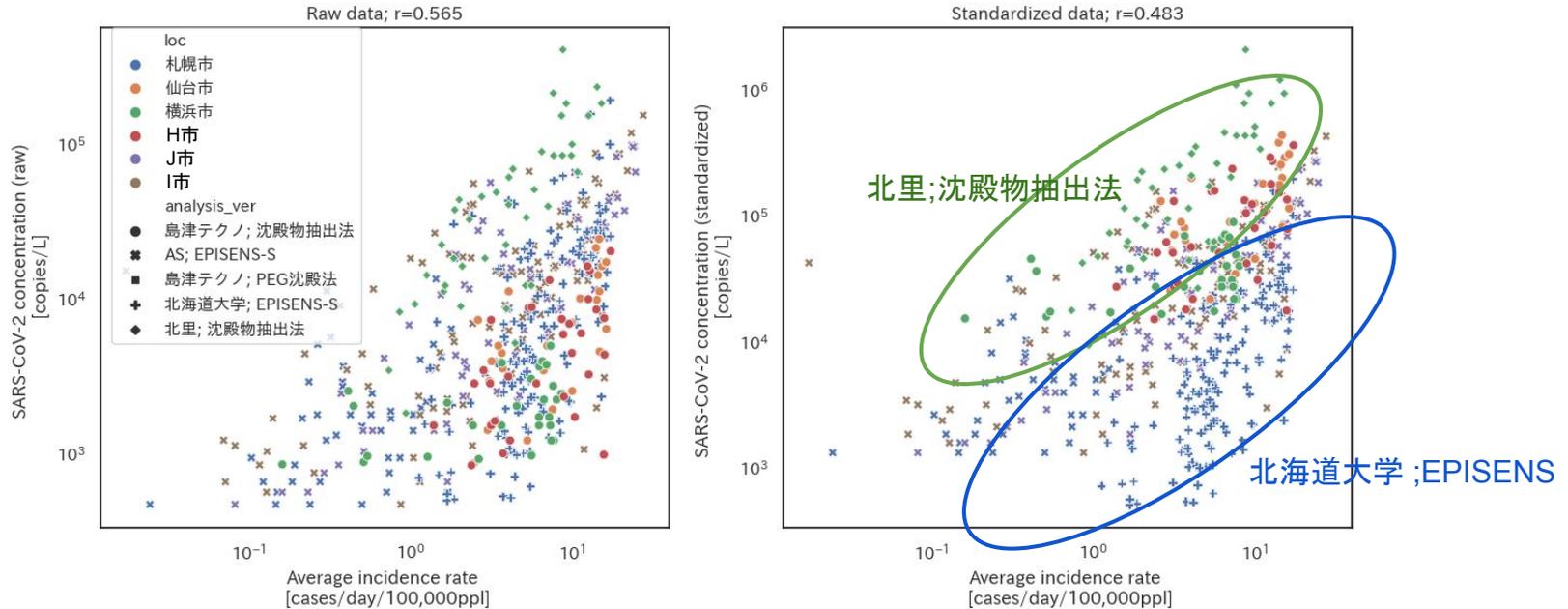


SARS-CoV-2 concentration (standardized)
[copies/L]



北里;沈殿物抽出法で濃度が非常に高く、北海道大学 ;EPISENSで濃度が非常に低く出
てしまっている。

フィールドデータとの比較：罹患率との比較（全手法）

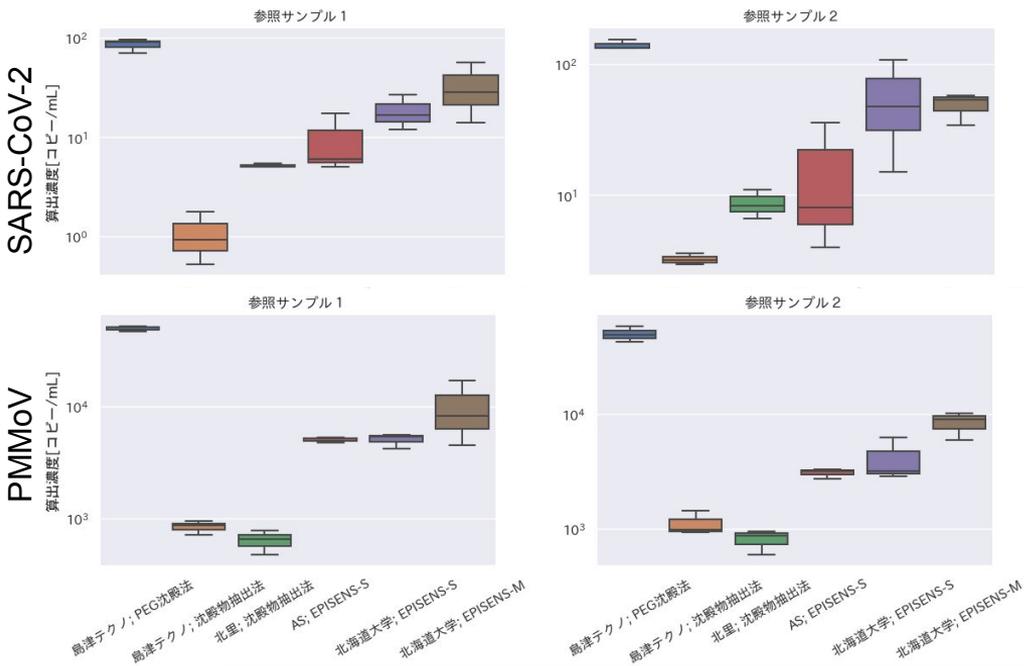


北里;沈殿物抽出法で濃度が非常に高く、北海道大学 ;EPISENSで濃度が非常に低く出
てしまっている。

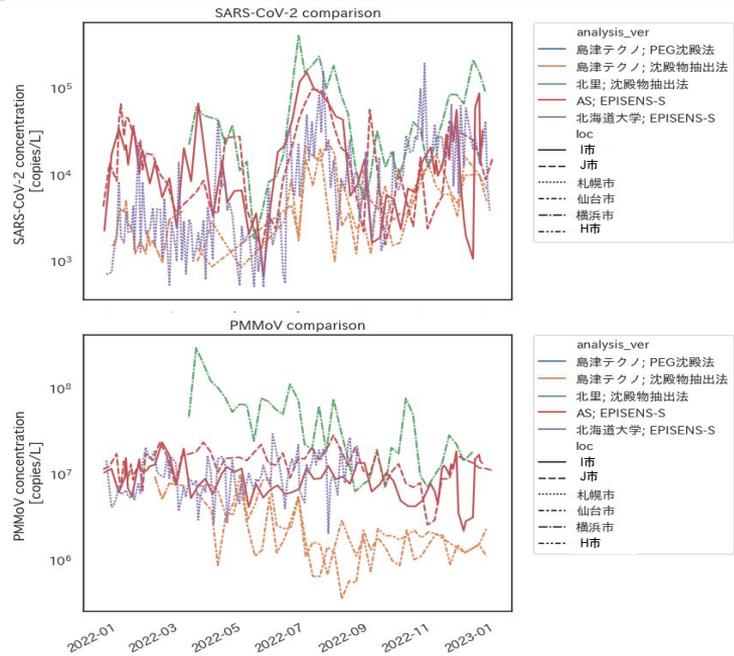


なぜ北里;沈殿物抽出法での標準化がうまくいかないのか？

標準化試験の結果



フィールドデータ

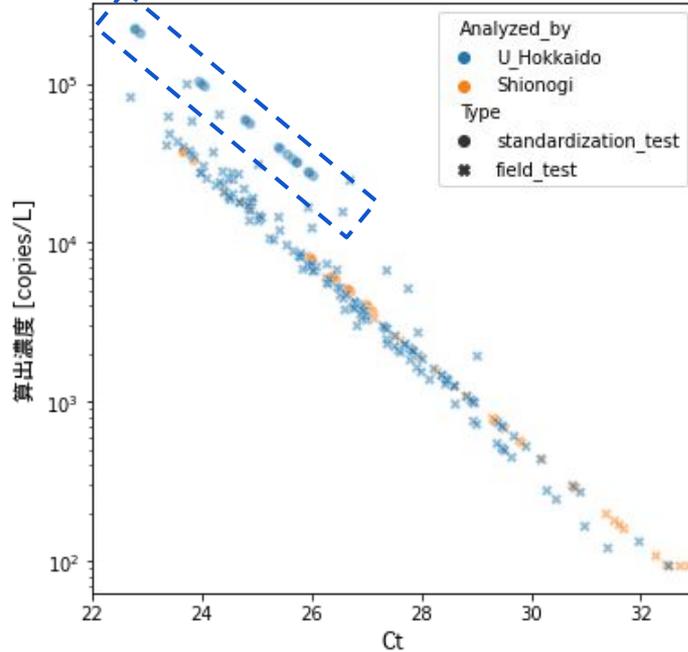


北里;沈殿物抽出法ではフィールド調査の結果が SARS-CoV-2、PMMoVともに思いのほか高い。→フィールド調査データの算出の仕方に問題がある可能性？

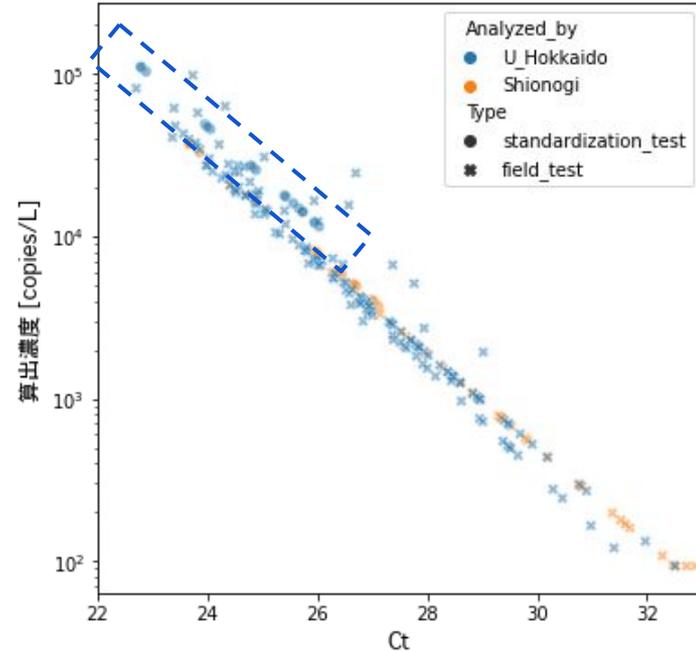


なぜ北大 ;EPISENSの標準化がうまくいかないのか？

訂正前



訂正後



標準化試験において、検量線に問題があった。修正後も、普段のデータとは Ct値から濃度算出する関係式が若干違うように見える。が、仮に普段の検量線を使ったとしても標準化は十分うまくいかない。うーむ。

成果のまとめ

- 分析手法・分析機関によって、同じサンプルを分析してもウイルス濃度測定値にかなりの違いが出る。
- ウイルス濃度をどれだけ過大過小評価するかは分析手法・分析機関ごとに特有であり、サンプルには依存しないと考えられることから、本試験で提案したような標準化パラメータを用いることにより分析手法・分析機関によるバラツキを補正できる可能性を示した(標準化)。
- 標準化により分析手法・分析機関間”のバラツキがある程度補正され、データの相互比較がしやすくなった。また、PMMoV補正等の正規化により分析手法・分析機関内”のバラツキも補正しうることが示された。標準化と正規化を合わせた規格化により、下水サーベイランスデータはさらに相互比較がしやすくなった。
- 分析手法は感度だけでなく、精度にも注意して検証する必要がある。分析手法によってはRNA濃縮・抽出時のウイルス回収にバラツキがあり、精度が高くないものもあった。
- 変異の出現などにより、特定のプライマーによる測定値に大きなバイアスがかかりうることがわかった。この事象は分析手法・分析機関にもよるが、こういった事象にまず気づき、対処するためにN1、N2プライマー両方をそれぞれ(少なくとも定期的に)定量化することが求められる。

今後の課題

- 今後下水サーベイランスデータをネットワークとして活用していく際に、本試験のようなデータ標準化試験を定期的実施していくことが求められる。その際、参照サンプルの準備、データの取りまとめ、試験実施の調整など行う試験実施の主体が産官学のどの組織にあるべきなのか議論する必要がある。
- 日本全体での標準化試験の定期的な実施が推奨される一方で、各分析機関も、手法の一部を変えた際(例:CRマスタミックスの変更)などには毎回、自身でデータの水平比較を可能にするような取り組みを行うことが推奨される。その際にも、本提案にある参照サンプルを用いた標準化(Fig.3)が活用できるため、各分析機関にもデータ標準化のノウハウの共有が重要である。
- 本試験では、分析機関内”のバラツキを補正するため、PMMoVを用いた勘弁なデータの正規化を試みた。データの正規化に関しては、PMMoV補正以外にも流量やその他環境因子を用いることも可能であり、今後さらに正規化手法に関する検討が必要である。
- 今後さらなる研究とデータの活用を促すために、国内の下水サーベイランスデータは一元化して格納することが理想である。そして、一元化して格納されたデータは、データの標準化・正規化・規格化されていると格段に利用のしやすさが上がる。異なる分析機関から算出されるデータを一元化して取り扱うためには、データアウトプットのフォーマットや、データを格納するデータベースを整える必要がある。