

厚生労働科学研究費補助金（食品の安全確保推進研究事業）
食品微生物試験法の国際調和に関する研究
平成 30 年度分担研究報告書

国際動向調査及び妥当性評価に関する研究

分担研究者 松岡 英明 東京農工大学 名誉教授

研究要旨：

食品微生物試験法を国際調和させるためには、国際動向を的確に把握し、わが国の妥当性評価ガイドライン案の更新に反映させることが必要である。そのため、ローザンヌ（スイス）で開催された ISO/TC34/SC9 の 2018 年次総会(6/19-6/22)に出席し、最新の国際動向を収集することとした。同会議では 25 のワーキンググループ（WG）、3 のアドホック WG の活動報告、及びそれらに関連する議論が行われた。懸案となっていた ISO16140:2016 版（代替法の妥当性評価に関するガイドライン）の内容で理解できなかった諸点については、その制定過程に携わっていた WG3（メソッドバリデーション）および WG2（統計学）の主旨等と直接議論した。特に、WG3 関連事項に関しては理解を深めることができた。最も重要な点は、代替法の方が、参照法よりも優れた性能を示す場合は、単に同等ではないとしてその代替法を棄却するのではなく、出来る限り、代替法を認証するようにしよう、との方針になっていることである。そして、そのために、念入りな手続きが新たに導入されていることを確認し、取り纏めた。一方、WG3 関連事項である統計学的評価基準については、WG2 の提言を自動的に採用しているとの情報を得た。WG2 の提言とは、数年にわたって集積された実験データに基づいて、経験的に定めた内容である、とのことであった。関連する WG に数年にわたって参加し議論を主導することは、現状では、困難と思われるので、少なくとも WG の主要メンバーと直接議論できるチャンネルを維持することが、今後のわが国における国際整合を踏まえた代替法の採用を検討するにあたっては重要な課題であると考えられる。

A．研究目的

食品微生物試験法を国際調和させるためには、我が国の試験法を ISO 16140 にしたがって妥当性確認する必要がある。そのため、当初は、ISO 16140:2003 に基づくガイドラインを作成した。しかし 2016 年に出された改訂版では内容が大幅に刷新されていた。そこで、改訂された内容を調査し、それを反映した新たなガイドライン作成に取り掛かった。

改訂版では、例えば、「ペアード(Paired)試験とアンペアード(Unpaired)試験」、「確定試験(Confirmation)」など、全く新しい概念の規定が加わっていて、その具体的な内

容の理解に苦慮した。この規格を扱っているのは ISO TC34/SC9 であるため、その 2017 年次総会(東京)に出席し、議論の経緯を調査した。しかし十分な理解は得られなかった。そこで、今年度 2018 年次総会(ローザンヌ、スイス)に出席し、継続調査することを目的とした。

本報告書では、こうして明らかになった内容をまとめた。

B．研究方法

ISO TC34/SC9 の 2019 年次総会に出席すると共に、ISO16140:2016 版の制定過程に携わっていた WG3（メソッドバリデーション

ョン)および WG2(統計学)の主査等と対面で議論するとともに、帰国後も必要に応じてメールでの意見交換を行った。

C. 研究結果

(1) ペアードスタディ (paired study) とアンペアードスタディ (unpaired study) の区別について

- ・スタディの最初の工程の増菌培養の条件が参照法と代替法が同じ場合はペアードスタディといい、条件が異なる場合をアンペアードスタディという。定性試験における感度評価で、「確認試験」の要求度が異なる。

(2) 定性試験における評価項目

- ・定性試験では次の3項目を実施する。すなわち、
自然汚染食品あるいは菌添加食品を用いて感度 (sensitivity) を明らかにする。この場合の菌濃度は陽性率 50% 程度 (25~75% の範囲)。菌添加食品を用いて菌濃度の検出の相対水準 (relative level of detection; RLOD) [適切に検出できる菌濃度] を求める。
代替法の包含性 (inclusivity) および排他性 (exclusivity) を求める。
- ・試料 (Sample) の数は 1 食品カテゴリーあたり、3 種類の食品タイプを選び、1 食品タイプあたり最低 20 個の試料数が必要。したがって、1 つの食品カテゴリーあたり、[食品タイプの数: 3 以上] × 20 以上 [試料 / 食品タイプ] = 60 以上。
- ・感度を明らかにする試験結果の整理と感度および関連指標 (相対真度、代替法擬陽性率) を求める計算式

	参照法陽性 (R ⁺)	参照法陰性 (R ⁻)
代替法陽性 (A ⁺)	陽性一致 PA	陽性偏差 PD

代替法陰性 (A ⁻)	陰性偏差 ND	陰性一致 NA
-------------------------	---------	---------

代替法の感度: $SE_A = (PA + PD) / (PA + ND + PD) \times 100\%$ 、

参照法の感度: $SE_R = (PA + ND) / (PA + ND + PD) \times 100\%$ 、

相対真度 (Relative trueness): $RT = (PA + NA) / N \times 100\%$ 、ただし $N = PA + NA + ND + PD$ ($N \geq 60$)、

代替法の擬陽性率: $FP = PD / N \times 100\%$ 。

- ・試験結果に基づく許容限界 (Acceptability limit; AL)

カテゴリー	ペアードスタディ		アンペアードスタディ
	ND - PD*	ND + PD	ND - PD*
1	3	6	3
2	4	8	4
3	5	10	5
4	5	12	5
5	5	14	5
6	6	16	6
7	6	18	7
8	6	20	7

*原文に記載はないが、絶対値と理解すべき。

(3) 新たに確認試験 (Confirmation) を追加する条件と同等性評価に及ぼす影響

- ・定性試験で、参照法と代替法の結果が異なった場合、従来は、単に陽性偏差、あるいは陰性偏差と結論。しかし、改訂版では新たに確認試験の実施が追加されている。ペアード試験では陽性偏差の場合のみ実施すべきとなっているが、アンペアード試験では全ての場合に実施するよう規定されている

る。

- 例えば、参照法(R)で陽性、代替法(A)で陰性のときは擬陰性となるが、引き続き実施した確認試験(C)が陽性か陰性かで結論が異なる。R、A、Cが陽性か陰性かで、以下の総計8ケースが考えられる。

- R: 陽性、A: 陰性→C: 陽性 擬陰性 (False negative) による陽性一致 (Positive agreement) :PA_{FN}*
- R: 陽性、A: 陰性→C: 陰性 陰性偏差 (Negative deviation) :ND
- R: 陰性、A: 陽性→C: 陽性 陽性偏差 (Positive deviation) :PD
- R: 陰性、A: 陽性→C: 陰性 擬陽性 (False positive) による陰性一致 (Negative agreement) :NA_{FP}
- R: 陽性、A: 陽性→C: 陽性 陽性一致 (Positive agreement) :PA
- R: 陽性、A: 陽性→C: 陰性 擬陽性 (False positive) による陰性偏差 (Negative deviation) :ND_{FP}
- R: 陰性、A: 陰性→C: 陽性 擬陰性 (False negative) による陽性偏差 (Positive deviation) :PD_{FN}*
- R: 陰性、A: 陰性→C: 陰性 陰性一致 (Negative agreement) :NA

なお、この中で*の表記は原文では誤記。

- 各ケースの数は次の表のようになる。

	参照法陽性 (R ⁺)	参照法陰性 (R ⁻)
代替法陽性 (A ⁺)	陽性一致 PA _T (PA+PA _{FN})	陽性偏差 PD _T (PD+PD _{FN})
代替法陰性 (A ⁻)	陰性偏差 ND _T (ND+ND _{FP})	陰性一致 NA _T (NA+NA _{FP})

T: 全体 (total)

代替法の感度: SE_A = (PA_T + PD_T) / (PA_T + ND_T + PD_T) × 100%

参照法の感度: SE_R = (PA_T + ND_T) / (PA_T +

$$ND_T + PD_T) \times 100\%$$

相対精確さ (Relative trueness): AC = (PA_T + NA_T) / N × 100% ただし N =

$$PA_T + NA_T + ND_T + PD_T \quad (N \geq 60)$$

代替法の擬陽性率: FP = (PD_T + ND_{FP} + NA_{FP}) / N × 100%

- この結果に基づく許容限界

カテゴリー	ペアードスタディ		アンペアードスタディ
	ND _T - PD _T *	ND _T + PD _T	ND _T - PD _T *
1	3	6	3
2	4	8	4
3	5	10	5
4	5	12	5
5	5	14	5
6	6	16	6
7	6	18	7
8	6	20	7

*原文では誤記。

- 具体的な事例は、添付資料1に掲載。

(4) 新たに確認試験を追加する理由

- ペアード試験では擬陽性結果の場合に必要とされ、擬陰性結果の場合必ずしも必要ではない。しかし、これは、擬陽性が擬陰性の結果よりも重大な問題だ、というわけではない。試験法は、標的菌をできるだけ高頻度に検出できる方が高性能と考える。代替法で検出できて、参照法で検出できないのは、代替法のほうが高性能であることを示唆している。そこで、その検出したものが確実に標的菌であることを確認するために確認試験(菌種の確認、同定)を行うようにしよう、との考えである。

- ・ 確認試験に用いる試験法にはいくつかの選択肢があり、その中には代替法の一部（菌種の確定試験）を利用する選択肢もある。その代替法全体としては、まだ、バリデーションされていないにもかかわらず、利用する菌種の確認試験の部分が他の試験法の中で、すでにバリデーションされている場合は、それを利用できる。

(5) RLOD を求める計算式

- ・ 定性試験では RLOD が要求されており、そのための EXCEL プログラムがオンラインで提供されている。便利ではあるが、計算の各段階の詳細はわからない。これに対して、

Annex D Models for RLOD calculations using data from the method comparison study. および

Annex F Considerations for calculations of the relative level of detection (RLOD) between laboratories as obtained in an interlaboratory study.

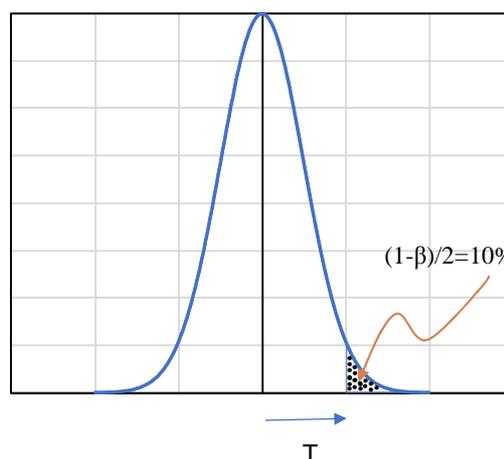
を参照して、各段階を手計算で行うことも可能、となっているが、理解困難な点は変わらない。

(6) 定量試験における許容区間

- ・ 定量試験では、代替法と参照法の試験結果から、その平均値の差（バイアス）と代替法のバラツキの大きさから、それが許容限界 $\pm[0.5 \times (\text{Log コロニー数})]$ の範囲内にあるかどうかによって、同等性を判定する。
- ・ 代替法のバラツキから、代替法の測定結果が 80% となるような菌数の範囲を求める。これが β -ETI (β -Expectation tolerance interval) である。 $t=0$ の両側に対称に広がる分布曲線 $f(t)$ で、 $t>T$ あるいは $-t<-T$ の値になる確率が α (例えば 0.05) 以下になるとき、 T の値を、両側検定による確率 α に対する t 値といって、有意差がある場合を

判定する t -検定で用いられる。同時に、 $f(t)$ が $-T<t<T$ の範囲にある確率は $1-\alpha$ であるといえる。

- ・ 同等性を判定する場合は β が用いられるが、その統計的処理は共通である。代替法の測定結果に基づく確率分布曲線として、 t -分布曲線を用いる。 $\beta=0.8$ となるような T 値は、両側検定による確率 $1-\beta=0.2$ として、 t -関数の逆関数 (EXCEL では $TINV$) を用いて求める (図 1)。これに、代替法の平均標準偏差 s_A と補正パラメータ $\sqrt{(1+1/n)}$ をかけた値が β -ETI/2 である。



T : Student-t分布の逆関数 (EXCELではTINV(両側検定1-β, dof)を用いて得られる値。

$$\beta\text{-ETI}/2 = T \times s_A \times \sqrt{1+1/n}$$

図1 . β -ETI(80%)説明図

- ・ 評価式は
 - $U_i = \text{バイアス(代替法の結果の平均値 - 参照法の結果の平均値)} + \beta\text{-ETI}(80\%)/2$
 - $L_i = \text{バイアス(代替法の結果の平均値 - 参照法の結果の平均値)} - [\beta\text{-ETI}(80\%)/2]$
 - $U_i < 0.5 \times (\text{Log コロニー数})$ 、 $L_i > -0.5 \times (\text{Log コロニー数})$ であれば、同等と判定する。
- ・ 図 2 の Accuracy profile の例では、6 種

類の菌濃度で、各濃度で5試料ずつ試験した結果を示している。この中で、2番目の菌濃度条件では L_i が範囲外になった。しかしこの場合、もし、 $4 \times$ (参照法の標準偏差 s_R)を超えなければ、依然として同等と判定する。

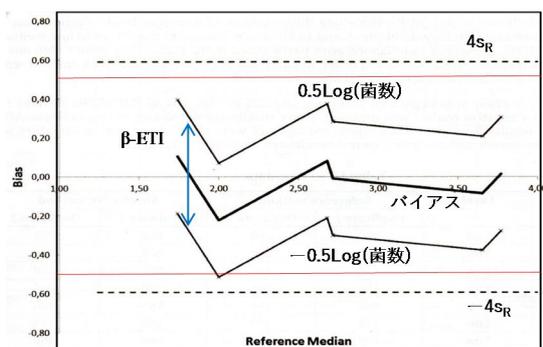


図2. Accuracy profile の例

- ここで問題となるのは、何故、85%、90%、あるいは95%ではなく80%なのか、また何故 $2 \times$ 、 $3 \times$ 、あるいは $5 \times$ ではなく $4 \times$ なのか？という疑問である。しかし、これに対しては、WG2「統計学」が議論して妥当な値だとして決めたこと、WG3は統計学の専門ではないので、WG2の提言をそのまま受け入れた、とのことであった。ただ、WG2としては、実験計画とデータの検証には数年を要したそうである。通常、議論された全ての内容が、最終的に出版されるドラフトに盛り込まれることはないが、何故、その内容だけが選択されたのか理由を理解することは難しい。
- なお、一旦、試験結果が許容範囲を外れたら、その根本原因を分析しなければならない。バリデーション・サーティフィケーション実施機関は、その分析結果に基づき、試験結果が受け入れられるかどうかの判断をする。こうした追加の調査は有用ではあるが、ISO16140-2にはその規定はない。

(7) その他

ISO 16140: 2016 で誤記と思われる内容を添付資料2にまとめた。

D. 結論

- ISO TC34/SC9のPメンバーとして年次総会に参加するようになって2年目の今年はWG2、WG3で規格策定に直接関わっている専門家と前回以上に緊密に議論できた。その結果、代替法のバリデーションの基本的なスタンスが理解できた。
- 第一に得た重要な認識は、参照法よりも代替法(新たに開発された方法)の方が、性能が優れている場合が多く、一概に「同等」ではないからといって、代替法を棄却することは不合理だ、との考えが強まっていることである。その代わりに、代替法の検出したものが、確かに標的菌であるということを、念には念を入れて確認する、ということである。バリデーションは規定通りに行えばよい、といった単純な考えは、もはや通用しない。試験法の本質を理解した人やチームによる、専門的な判断が常に要求されると考えるべきである。
- それゆえ、バリデーションのガイドラインの中身を議論するほどに、益々、我が国にもAOAC、AFNOR、NordVal、などに相当する実施監督機関が不可欠であるのと思いが強まった。
- なお、今のところ代替法は培養法に限定されている。非培養法に基づく代替法は、まだ同等以上として認証された例はない。数年前にCENから提案されたフローサイトメトリー法による生菌死菌計数法は、その後の議論の展開は見られない。非培養法に対しては、やはり、判断が難しく、最終的には生菌標準物質が不

可欠ではないかと推察される。

E . 健康危害情報

該当なし。

F . 研究発表

- (1) 国内合計 0 件
- (2) 海外合計 0 件
- (3) 学会発表 合計 1 件
 - 1 . 松岡英明：バイオにおける確からしさと不確かさ. (「電気化学と生命科学」企画シンポジウムにおける依頼講演)、電気化学会、平成 31 年 3 月 27 日、京都大学
- (4) 研究班会議：バリデーション部分会議等：合計 1 回
 - 1 . 第 36 回バリデーション作業部会、平成 30 年 11 月 28 日(水)、バイオメリユー・ジャパン(株)会議室、参加者 10 名、NIHSJ-2(カンピロバクター定性法)の培地組成変更について、他 5 課題について議論)

G . 知的所有権の取得状況

該当なし。

AFNOR 認証システムで、確認試験を行った例

Category	Type	PA	NA	PD	ND	PPND	PPNA	SE _{PA} %	SE _{NA} %	RT %	FPR %
1	a Ready to eat	7	10	3	1	0	0	90.9	72.7	81.0	0.0
	b Ready to reheat	7	10	2	1	0	0	90.0	80.0	85.0	0.0
	c Pastries, desserts, omelettes	8	12	0	1	0	0	88.9	100.0	95.2	0.0
	Total	22	32	5	3	0	0	90.0	83.3	87.1	0.0
2	a Raw, frozen, seasoned	10	10	3	4	0	0	76.5	82.4	74.1	0.0
	b Ready to eat and processed meat products	3	14	2	2	0	0	71.4	71.4	81.0	0.0
	c Delicatessen	4	13	3	3	0	0	70.0	70.0	73.9	0.0
	Total	17	37	8	9	0	0	73.5	76.5	76.1	0.0
3	a Raw dairy products	10	10	3	0	0	0	100.0	76.9	87.0	0.0
	b Pasteurized milk cheese	3	11	3	3	0	0	66.7	66.7	70.0	0.0
	c Ice cream, milk (pasteurized), flavored milk	5	12	2	1	0	0	87.5	75.0	85.0	0.0
	Total	18	33	8	4	0	0	86.7	73.3	81.0	0.0
4	a Raw vegetable products (fresh, frozen)	8	16	1	1	0	0	90.0	90.0	92.3	0.0
	b Mapped vegetables and heat processed vegetables	7	6	4	2	1	0	78.6	71.4	65.0	16.7
	c Vegetables based preparations, processed vegetables	4	15	1	2	0	0	71.4	85.7	86.4	0.0
	Total	19	37	6	5	1	0	80.6	80.6	82.4	2.7
5	a Raw products (fresh, frozen)	8	10	1	1	0	0	90.0	90.0	90.0	0.0
	b Smoked, marinated	7	10	3	0	0	0	100.0	70.0	85.0	0.0
	c Ready to eat or ready to reheat	7	12	0	3	0	0	70.0	100.0	86.4	0.0
	Total	22	32	4	4	0	0	86.7	86.7	87.1	0.0
6	a Process water	7	14	0	0	0	0	100.0	100.0	100.0	0.0
	b Dusts	5	11	4	5	0	1	64.3	71.4	65.4	9.1
	c Wipes	9	15	0	2	0	1	81.8	100.0	92.6	6.7
	Total	21	40	4	7	0	2	78.1	87.5	85.1	4.8
All categories		119	211	35	32	1	2	82.4	81.3	83.0	1.4

前増菌培養法が異なるために、アンペアードスタディとなり、 $|ND_T - PD_T| = |32 - 35| = 3$ 。6 カテゴリーに対する AL は 6 であるから、この場合は同等と判断される。

(2) 3M : MDA *Cronobacter* spp. の事例

Category	Type	Protocol	Study design	PD	ND	PPND	Unpaired study design		Paired study design			Combined unpaired and paired		
							(ND+PPND)-PD	AL	(ND+PPND)-PD	AL	(ND+PPND)+PD	AL	(ND+PPND)-PD	AL
1	a	General	Paired	0	0	0								
		Specific 1	Unpaired	2	1	0	-1							
	b	General	Paired	0	1	0		1		1				
		Specific 1	Unpaired	4	2	0	-2							
	c	General	Paired	0	0	0		0		0				
		Specific 1	Unpaired	4	1	0	-3							
Total 10g				0	1	0		1	3	1		6		
Total 300g Protocol 1				10	4	0	-6	3						
2	a	General	Paired	0	0	0								
		Specific 2	Unpaired	2	1	1	0							
	b	General	Paired	0	0	0		0		0				
		Specific 2	Unpaired	1	4	0	3							
	Total 10g				0	0	0		0	3	0		6	
	Total 300g Protocol 2				3	5	1	3	3					
3	a	General	Paired	0	1	0			1		1			
		General	Paired	0	2	0			2		2			
	b	General	Paired	0	0	0			0		0			
		General	Paired	0	0	0			0		0			
	Total				0	3	0		1	3	1		6	
	General Protocol (total paired)				0	4	0		4	5	4		19	
Specific Protocol 1				10	4	0	-6	3						
Specific Protocol 2				3	5	1								
Specific protocols (total unpaired)				13	9	1	-3							
Total				13	13	1		4	1	1	1	1	5	

赤枠はアンペアードスタディで、 $|ND_T - PD_T| = 3$ 。6 カテゴリーに対する AL は 6 であるから、この場合は同等と判断される。

青枠はペアードスタディで、 $|ND_T - PD_T| = 4$ (6 以下)、 $ND_T + PD_T = 4$ (16 以下)、あるから、この場合は同等と判断される。

ISO 16140-2: 2016 の誤記と史料された内容に係る確認状況

ISO 16140-2:2016 について、WG 議長へ以下の内容について確認を行った。

- (1) 確認試験の結果に基づき、擬陽性か擬陰性かの判断の表記に誤りがあると思われる。
該当箇所は、[5.1.3.4 Calculation and interpretation for sensitivity] (定性試験のシングルラボスタディにおける) の Table 1-4、[5.2.3 Calculations and summary of data] (定性試験のコラボスタディにおける) の Table 9-11 である。
- (2) 定性試験における許容限界を示す表で、(ND-PD)が絶対値になっていない。
- (3) [5.2.3 Calculations and summary of data] (定性試験のコラボスタディにおける) で、式(6)(7)に記載されている P_0 、 CP_0 の説明は誤解されるかも知れない。ここでは、参照法では擬陽性の総数 P_0 、代替法では確認試験後の擬陽性総数 CP_0 を用いる、と明記すればよいと思われる。ただし、代替法では、必ずしも確認試験をするとは限らないので、そのことを付記すべきであるが、その記載がない。
- (4) [5.2.3 Calculations and summary of data] (定性試験のコラボスタディにおける) で、「 N_X in equations (12)-(14)」は「 $N_{X(R+A)}$ for equation (14) is the total number of N_{XR} and N_{XA} 」のように参照法と代替法の場合を区別した表記にすべきと思われる。
- (5) [6.2.3 Calculation, summary, and interpretation of data] (定量試験のコラボスタディにおける)
Table 18 の Suffixes of collaborator (i), level (k), and replicate (l) in Table 18 は collaborator (k), level (i), and replicate (j) のように本文と合わせるようにすべきと思われる。
- (6) [Annex H Table H.1]
Calculated results of Step 1 (X_i) and Step 2 (Y_i) における以下の記載について、
 X_i (1.740, 2.114, 2.681, 2.716, 3.653, 3/771) は正しくは(1.780, 2.179, 2.657, 2.682, 3.685, 3.778)
 Y_i (1.845, 1.778, 2.763, 2.708, 3.568, 3.785) は正しくは(1.181, 1.728, 2.800, 2.700, 3.666, 3.839)
ではないか。これらは簡単な平均値の計算での間違いであり、考えが正しければ、後続の計算結果 Step3 ~ Step7 全ても誤記となるのではないか。
- (7) 以上の意見に対し、
『残念ながら、指摘された内容を直ちに改訂することはできないため、次回の改版時、あるいは修正時に検討されるようにしたい。その際には、提出できるよう準備をお願いしたい。』との回答を得た。