

Validation of Performance in Library-Dependent Microbial Source Tracking Studies

Valerie J. Harwood, Ph.D.

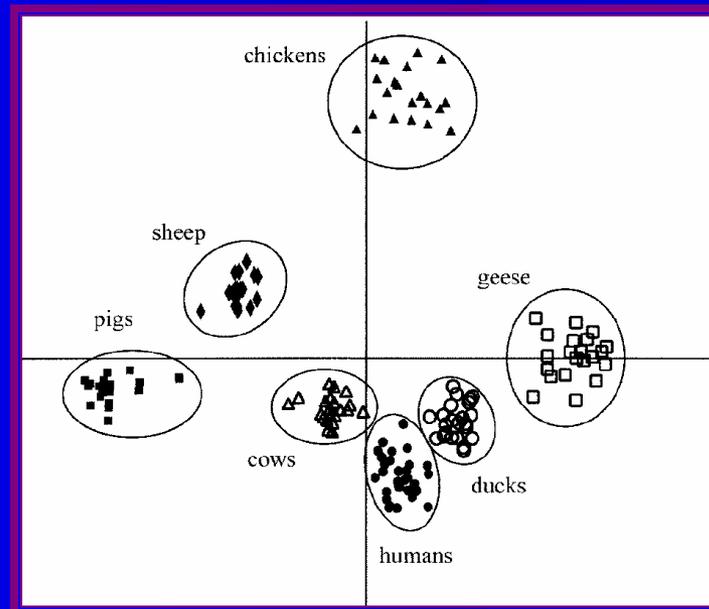
Department of Biology

University of South Florida



What Is A Library?

- Isolates cultured from fecal material whose host source is **KNOWN**
- Phenotypic or genotypic typing discriminates among subtypes



Dombek et al 2000
AEM 66:2572

How Is a Library Used?

- Isolates are typed and “matched” to subtypes in library.

- The “match” can be statistical

- discriminant analysis
- K-nearest neighbor
- average similarity

- or 1-to-1

- maximum similarity
- “exact matching”

Site	Human n (%)	Sediment n (%)	Birds n (%)	Total n (%)
MS	18 (34.0)	17 (32.0)	18 (34.0)	53 (100)
MAR	11 (21.1)	34 (65.4)	7 (13.5)	52 (100)
BR	12 (20.0)	31 (51.7)	17 (28.3)	60 (100)
319	23 (29.9)	36 (46.7)	18 (23.4)	77 (100)

Based on Early MST Studies, Expectations of the Methods Reflected

- Wild optimism



Caveats

- **Libraries were generally small (hundreds of isolates).**
- **Libraries retained “identical” subtypes from fecal samples.**
- **Validation was solely by rate of correct classification (RCC).**
- **This is an internal validation measure that does NOT test the library’s ability to classify isolates from fecal material that is not part of the library.**

A Second Generation of Studies Suggested:

Uh-oh...not so fast!

“Proficiency” isolates were not accurately classified.

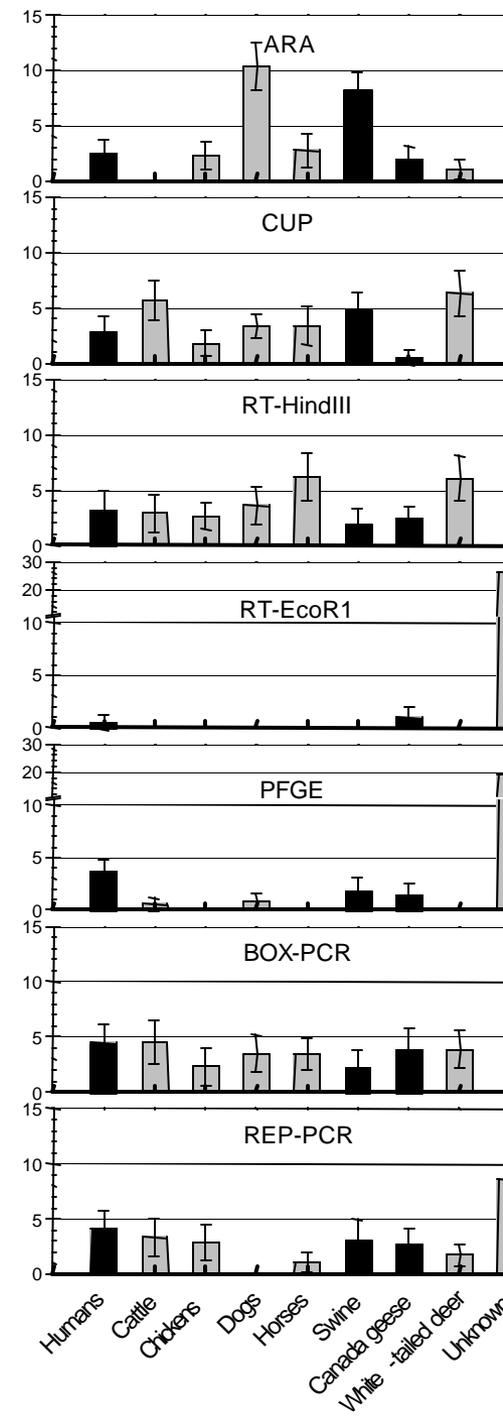
SCCWRP study 2003

Stoeckel et al 2004

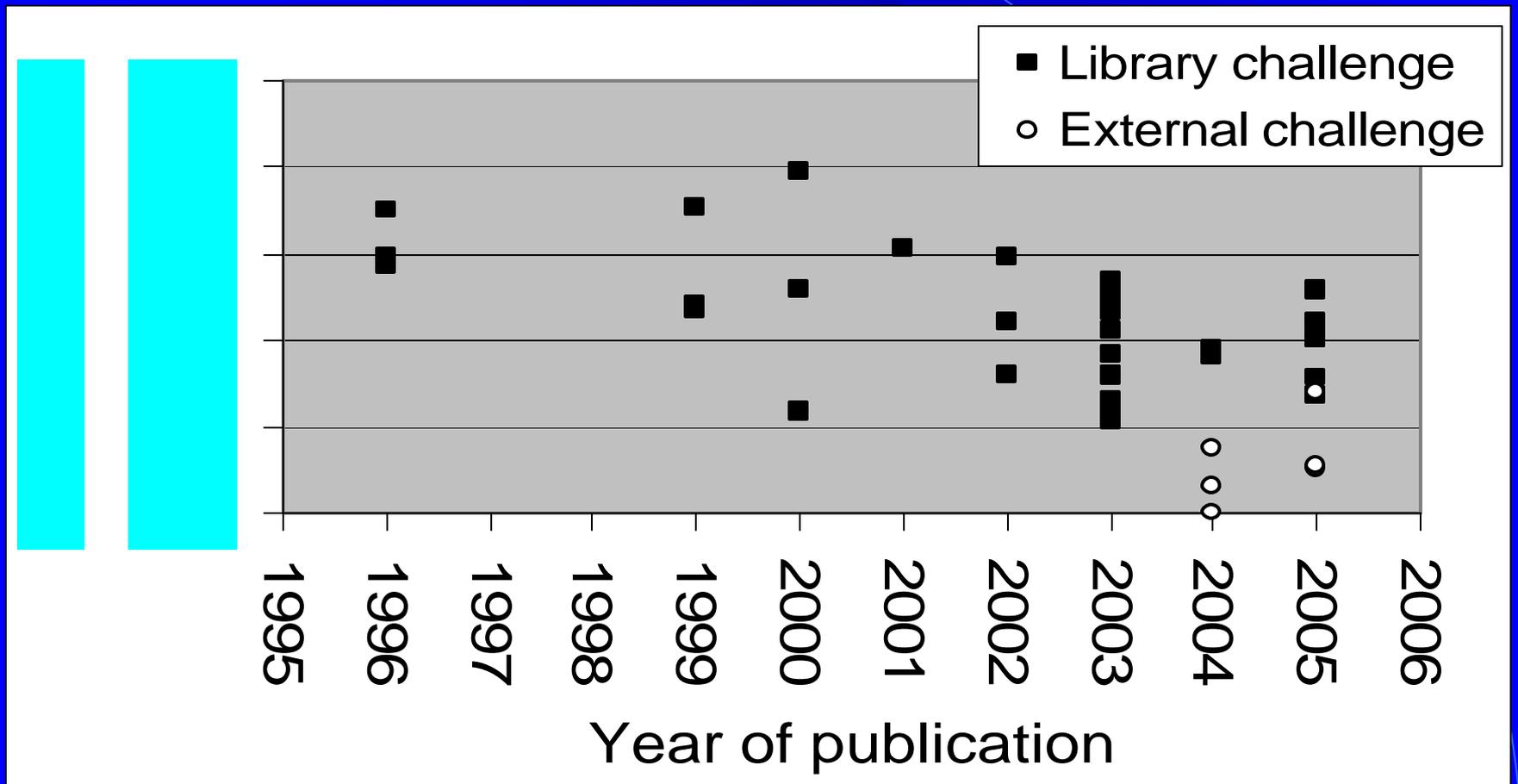


Moore et al 2005

30 *E. coli* isolates were chosen randomly from the challenge sample set
10 human
10 swine
10 Canada goose



Accuracy Trend for Library-Based Methods in Published MST Studies



Validation Strategies

Internal Validation of Library RCC by Jackknife

- **Hold-one-out**
- **Pulled sample**

External Validation of Library Accuracy: Challenge (Proficiency) Tests

- **Blind isolates from known sources**
- **Seeded samples**
- **Environmental waters contaminated from a known source**

Performance Measures

- **Sensitivity (rate of correct classification; ability to detect source when present)**
- **Specificity (discriminatory capability of method; ability to rule out source when absent)**

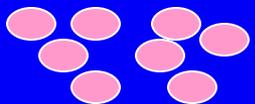
Sensitivity

Ability to detect target when present = RCC

% of actual + that are detected



10 samples contain “human” source, but only 8 test positive

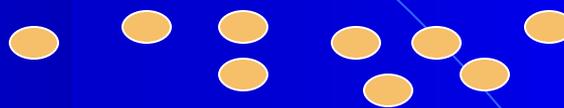


$$8/10 = 80\%$$

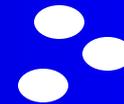
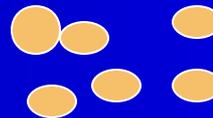
Specificity

Ability to discriminate among sources

Proportion of negative samples that test negative



9 samples contain no “human” source, but only 6 test negative



False +

$$6/9 = 66.7\%$$

Comparing Apples to Oranges - How to Compare Method Accuracy When the Possible Number of Source Categories is Different?

Example: Study A splits all observations into two possible source categories, e.g. animal and human, and the method correctly assesses fecal source in 74% of samples.

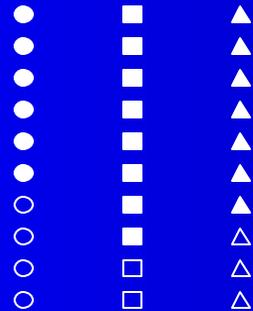
Study B splits all observations into four possible source categories, and the fecal source is assessed correctly in 55% of samples.

“Benefit Over Random”

2-Category Split

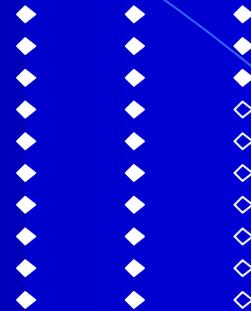


Nonhuman
source isolates



Correct: 210 of 300
RCC: 70%

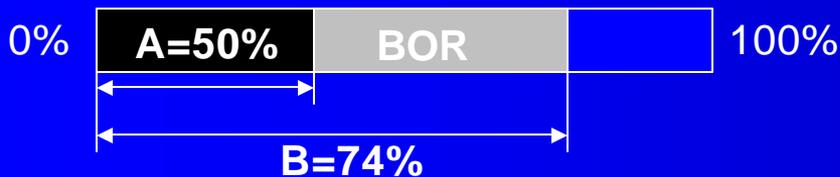
Human-source
isolates



Correct: 230 of 300
RCC: 77%



Classification accuracy



A=measure of random classification (e.g. $1/k$)
B=measure of accuracy (e.g. ARCC)

Benefit over random (BOR) = B - A

ARCC: 74%
Categories: Two
Random: 50%
BOR: 74%-50%=24%

4-Category Split



Correct:
RCC:

50 of 100
50%

80 of 100
80%

30 of 100
30%

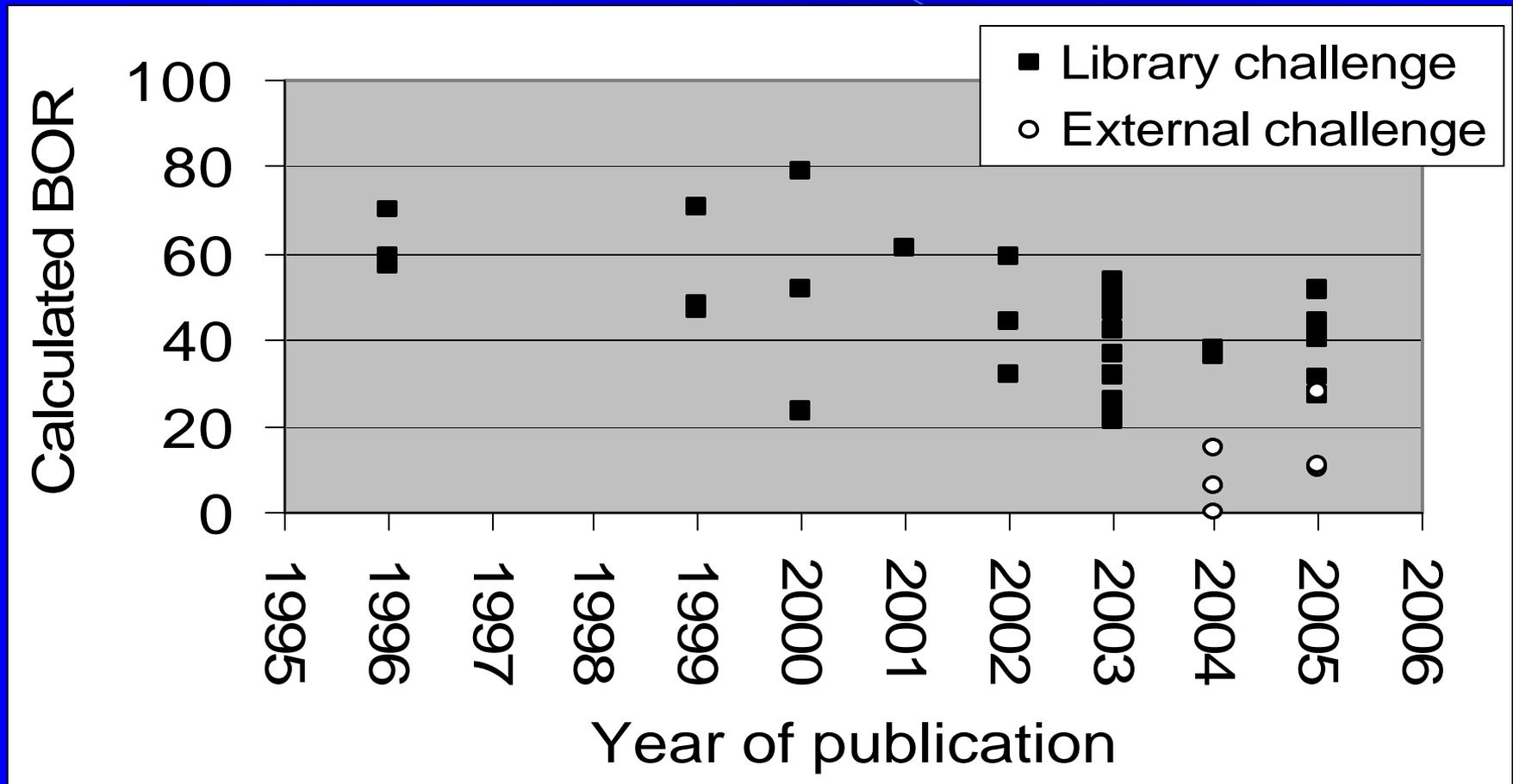
60 of 100
60%

Each symbol represents ten isolates. Dark symbols were correctly classified Open symbols were incorrectly classified

ARCC: 55%
Categories: Four
Random: 25%
BOR: 55% - 25% = 30%

By comparison, the two-way split had ARCC 74%, BOR 24%

BOR of Library-Based Methods



Building a Better Library

- **Use common sense and local knowledge in selecting source categories**
- **Sample many individuals of each host species**
- **Estimate indicator organism diversity before typing**
 - **Composite fecal samples will contain a much higher-diversity microbial population than individual samples**
- **Declone at the sample level**

**External Measures of Method
Success Should Be REQUIRED
in Publications and for
Management Reports
(Defensibility)**



Field Validation Needed

- Confidence estimates for classification of unknowns
- Effects of differential survival/ rapid die-off in secondary habitat
- Matrix effects such as humic substances on PCR (LIMs)



**Confirm Successful
Methodology Transfer!**

Case Study 1: Wakulla County Florida BOX-PCR of *Enterococcus* spp.

Source categories

- **Human (septic pump-out and WTP influent)**
- **Bird (gulls and pelicans)**
- **Sediment**

Library

- **Initially 778 isolates**

Composition of the *Enterococcus* BOX-PCR Library

Source	Isolates typed	Isolates in Decloned Library	Number of samples
Lift stations	212	57	5
Wastewater influent	105	45	2
Septic tank	48	12	2
Bird feces	150	55	41
Marine sediment	263	121	26
Total	778	290	76

Internal Library Accuracy (Sensitivity) by Jackknife Analysis

Assigned Source Categories

True Source	Human (%)	Marine Sediment (%)	Birds (%)	Total (%)
Human ^a	<u>61.4</u>	28.9	9.7	100.0
Sediment	28.5	<u>57.0</u>	14.5	100.0
Birds	27.3	12.7	<u>60.0</u>	100.0

BOR = 27%

Proficiency Isolates

Assigned Source Categories

True Source	Human n (%)	Sediment n (%)	Birds n (%)	Total n (%)
Human	<u>88 (61.1)</u>	30 (20.8)	26 (18.1)	144 (100.0)
Sediment	14 (31.1)	<u>26 (57.8)</u>	5 (11.1)	45 (100.0)
Gulls	3 (33.3)	1 (11.1)	<u>5 (55.6)</u>	9 (100.0)
<u>Ducks</u>	32 (31.1)	41 (39.8)	<u>27 (26.2)</u>	103 (100.0)

The library contained no duck isolates



Classification of Isolates from Water by Bootstrap Analysis

Site	Human n (%)	Sediment n (%)	Birds n (%)	Total n (%)
Beach	18 (34.0)	17 (32.0)	18 (34.0)	53 (100.0)
Marsh	11 (21.1)	34 (65.4)	7 (13.5)	52 (100.0)
Boat ramp	12 (20.0)	31 (51.7)	17 (28.3)	60 (100.0)
Bridge	23 (29.9)	36 (46.7)	18 (23.4)	77 (100.0)

Water Isolates Classified with = 80% Bootstrap Value

Site	Human n (%)	Sediment n (%)	Birds n (%)	Total n (%)
Beach	1 (16.7)	1 (16.7)	4 (66.6)	6 (100.0)
Marsh	0 (0.0)	1 (100.0)	0 (0.0)	1 (100.0)
Boat ramp	1 (7.6)	6 (46.2)	6 (46.2)	13 (100.0)
Bridge	3 (17.6)	6 (35.3)	8 (47.1)	17 (100.0)



Case Study 2: The Beach

Chuck Hagedorn, VA Tech

Investigation of high *Enterococcus* levels at beaches

- *Enterococcus* BOX-PCR library of 1681 isolates
- *Enterococcus* ARA library of 2360 isolates
- Library isolates collected over a 2.5-year period

Blind proficiency samples were seeded with

Enterococcus isolates from various sources

Categorization of Isolates in Samples into One of Six Categories

Human only	4 correct of 4 (100%)
Dog only	2 correct of 4 (50%)
Gull only	3 correct of 4 (75%)
Animal dominant	3 correct of 4 (75%)
Human 50:50	3 correct of 4 (75%)
Human dominant	3 correct of 4 (75%)
Totals	18 correct of 24 (75%)
BOR	75% - 16.7% = 58.3%

The Smoke Test



Expectations of MST Stage 3

“Optimistic skepticism” Stoeckel 2006

- **Assess sensitivity and specificity**
- **Libraries must be validated by challenge with isolates and/or samples from independent reference materials (e.g. fecal samples)**
- **Library-independent methods must be validated by composite samples containing fecal material from target or nontarget sources**

Use Libraries In Appropriate Circumstances

- **Where \$ and incentive exist to maintain and update libraries.**
- **Where possible contamination sources are limited.**
- **In conjunction with library-independent methods.**

A photograph of a sunset over the ocean. The sun is low on the horizon, casting a golden glow across the sky and reflecting on the water. The sky is filled with soft, orange and yellow clouds. In the foreground, a dark silhouette of a person is visible on a beach, looking out at the sea.

Questions?

Contact:
vharwood@cas.usf.edu

