HOW TO WRITE A RESEARCH PAPER & STAY SANE

Agatha N. Jassem, PhD, (D)ABMM, FCCM
AMMI - CACMID Annual Conference, May 3 2017
Clinical Microbiologist, BCCDC Public Health Lab
Clinical Assistant Professor, U. of British Columbia
agatha.jassem@bccdc.ca

Conflict of Interest

I do not have an affiliation (financial or otherwise) with a pharmaceutical, medical device or communications organization

Objectives

- Describe how to focus a manuscript by including and directing text based on impact
- Format a manuscript by understanding paper structures and standards
- Recall how to write clearly and efficiently based on tips and established techniques

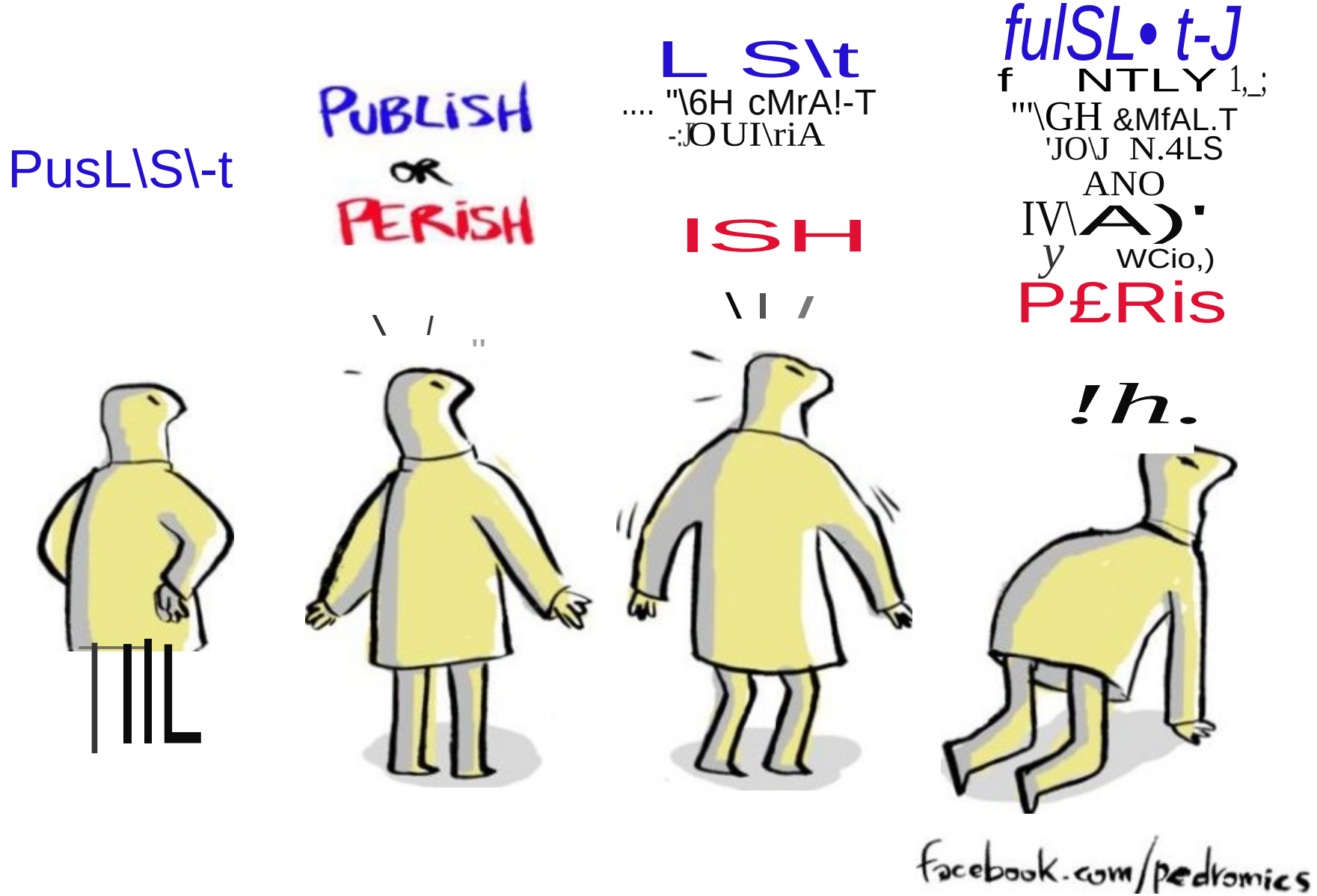
Why publish?

- communicate findings
- for review & being part of scientific community
- others can confirm/expand on findings
- others can apply findings
- need less publication bias

"If I have seen further than others, it is by standing on the shoulders of giants"

- Sir Isaac Newton

- for funders/employers



When? Now!

- Forces clarity and focus
- Makes for better research
- Useful for grants, talks, etc
- Start with an outline

How? Focus, Focus, Focus

What is the idea/contribution?

Need to convey something useful

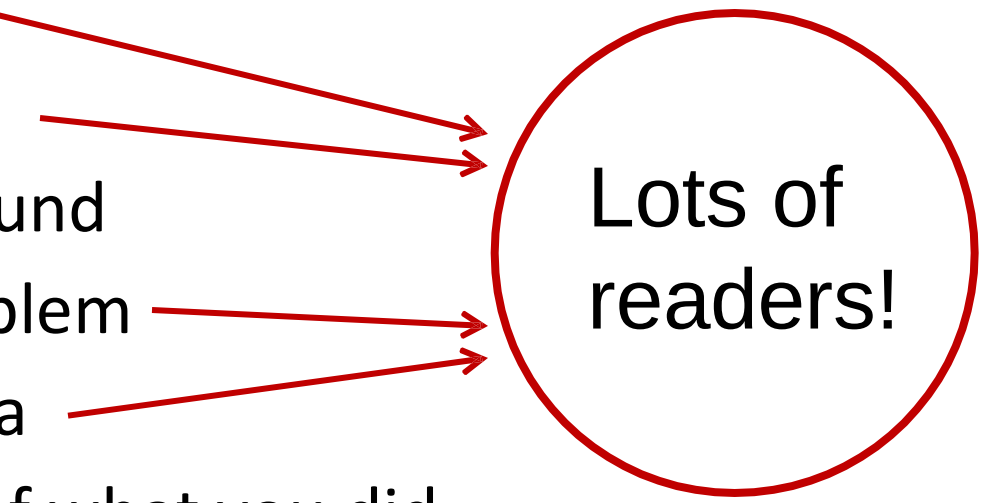
>1 idea = consider >1 paper

"In writing, you must kill all your darlings"

- William Faulkner

Format

- Title
- Abstract
- Background
- The problem
- Your idea
- Details of what you did
- Details of what you found
- How your work compares to related work
- Limitations of your work
- Conclusions
- Thanks!



Format in a writing timeline

- Outline + abstract?
- Details of what you did (methods)
- Details of what you found (results)
- Background + problem + idea (introduction)
- Related work + limitations (discussion)
- Conclusions + acknowledgements
- Abstract
- Title

Title

What did you do vs **what did you find**

Diagnostic performance of laboratory methods during a mumps outbreak in British Columbia

vs

Buccal swab RT-PCR provides highest diagnostic yield during an outbreak of mumps in a partially vaccinated population

Abstract

- Brief summary of each of the main sections
- Include only key points/information

Do: be descriptive

Don’t: include references

Why are you doing this?

What did you do?

What did you find?

What does it mean?

What is it good for?

Introduction

Do: present an interesting, unsolved problem

Do: state the method of investigation

Do: describe your idea/contribution

- claims should be refutable

Don't: provide a paper outline

Don't: use too much expert jargon

Don't: write a review

- limit related work to highlight major contributions, give examples of the problem, or justify your idea

Shiga toxin-producing *Escherichia coli* (STEC) is a major cause of sporadic and outbreak-associated enteric illness worldwide. While more than 100 STEC serotypes can cause illness in humans (1), traditional testing methods focus primarily on the most common serotype, O157:H7, owing to the ease of its detection using culture-based media. These methods underdetect non-O157 serogroups, and as such, their clinical burden is not well understood (2–4). Moreover, recent evidence suggests that infections attributed to non-O157 STEC may be more prevalent than those attributed to O157 (1, 5), with common serogroups being O26, O45, O103, O111, O121, and O145 in the United States (6, 7) and in other countries (1). Although non-O157 STEC serotypes are generally associated with milder disease than O157 STEC, infections caused by non-O157 STEC can also lead to hemolytic uremic syndrome (1) and have been associated with major outbreaks, most notably a 2011 outbreak in Germany that involved 3,816 cases and 54 deaths (8). Enhanced surveillance evidence is needed to determine the true burden of non-O157 STEC for public health investigations of potential exposure sources. Such efforts are ongoing in Canada (9) but require improved and comprehensive screening methods.

The Centers for Disease Control and Prevention (CDC) recommend that, in addition to O157-selective culture-based screening, all stools submitted for testing from patients with acute community-acquired diarrhea be assayed for non-O157 STEC with a test that detects the Shiga toxins or their genetic determinants (stx_1 and stx_2) (10). Given the low prevalence of STEC infections (1 to 2%) (11–13), the associated costs for clinical laboratories for personnel, equipment, and reagents are major barriers for implementation of the recommended universal screening of stool for STEC (4, 14). To reduce testing costs, we evaluated the feasibility of a pooled nucleic acid amplification test (NAAT) as an approach for low-cost high-throughput screening for stx_1 and stx_2 from stool. Similar approaches have been successfully applied for the detection of HIV (15, 16), hepatitis C (16, 17), and malaria (18, 19). Pooled NAAT strategies are best suited for scenarios where testing volumes are high but disease prevalence is low, as would be the case for universal stool screening.

Where is problem example?

Where is rationale?

What is the idea?

Is the contribution refutable?

Where is the related work?

Recently, European hospitals reported changing patterns in the molecular epidemiology of CDI, suggesting that the circulation of BI/NAP1/027 strains in some areas is decreasing while other strain types are emerging [22–24]. Furthermore, the incidence of community-associated (CA) CDI is increasing across Europe and North America [25–27] and occurring in patients that are younger and healthier and have fewer of the risk factors associated with hospital-associated (HA) CDI [28–32]. The current lack of comprehensive surveillance data impedes our ability to detect changes in the molecular epidemiology of CDI [25]. Characterization of CDI in nonoutbreak settings to include CA- and HA-CDI cases would allow for a comprehensive understanding of the pathogenicity and associated markers of strains responsible for CDI and also enable the detection of early shifts in molecular epidemiology that indicate the emergence of novel outbreak strains. In the present study, we determined the molecular epidemiology of *C. difficile* infections in British Columbia during a month-long, province-wide study in 2008 that involved hospital and community specimen collection sites. A follow-up study with one of the hospital sites was carried out in 2013 to determine distribution of strain types between the two time points and between CA- and HA-CDI cases. Isolates were characterized by PFGE, PCR-ribotyping, their carriage of toxin genes and deletions in *tcdC*, and susceptibilities to antimicrobials.

A lot of related work in two sentences

(to guide later discussion)

Methods used

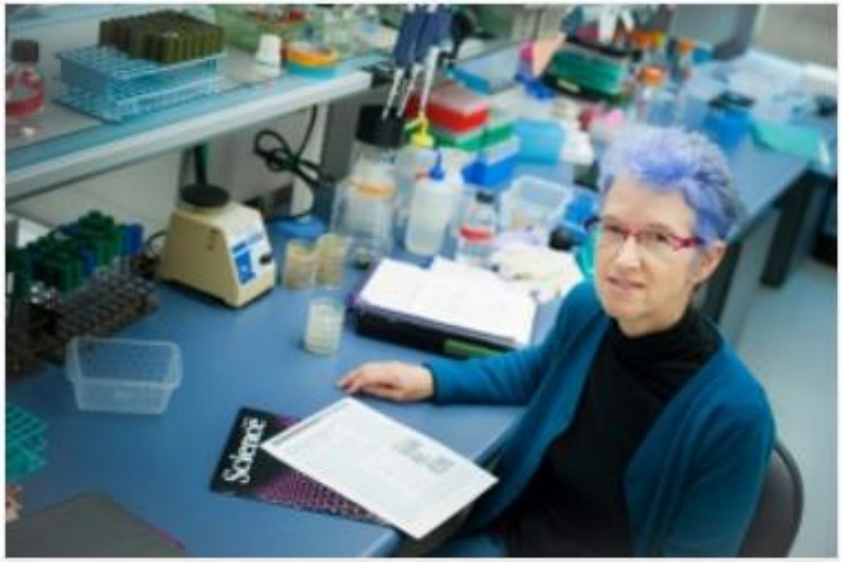
NATURE | NEWS

Study challenges existence of arsenic-based life

Open-science advocates fail to reproduce controversial findings.

Erika Check Hayden

20 January 2012



M. DEE/NATURE

Rosie Redfield says her findings are a "clear refutation" that the bacterium GFAJ-1 can incorporate arsenic into its DNA.

Methods

Purpose:

research can be reproduced

Materials & Methods

- Include exact technical specifications and quantities and source or method of preparation
- Use of generic or chemical names is preferred
- Organisms/subjects/samples should be identified accurately with sources/characteristics
  - put in a table if lots here
  - add selection criteria
  - add ethics approval info where needed

2.2. Control samples

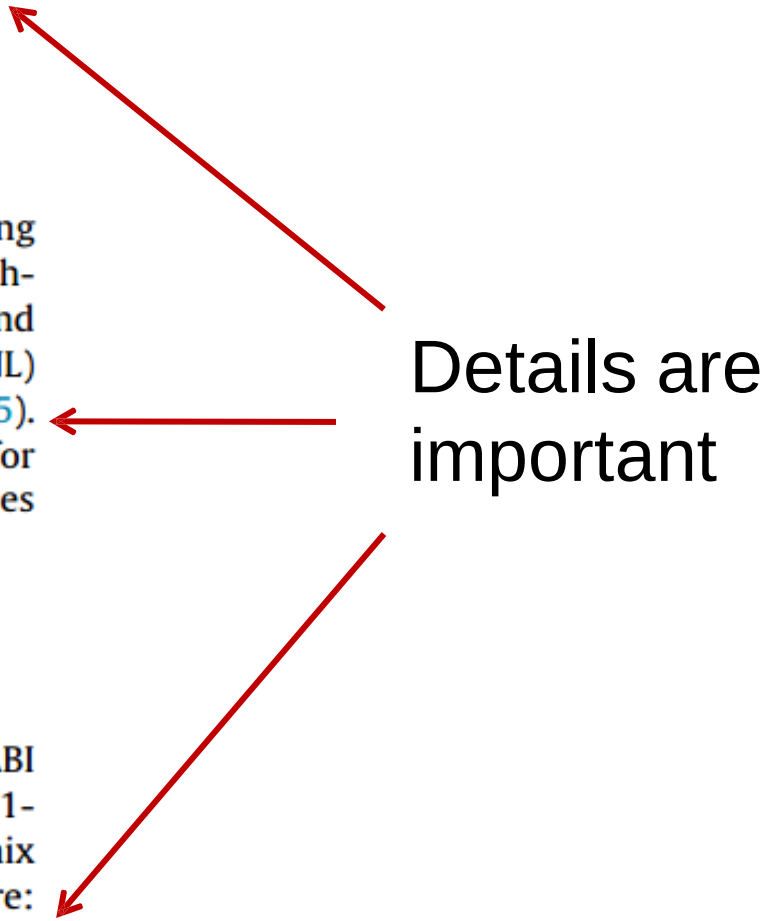
Custom-designed gBlock oligomers representing HCV Gt 1a, 1b, 2, and 3a (Integrated DNA Technologies (IDT)) diluted in HCV RNA negative extract were used as positive controls in each RT-PCR run. Additional positive controls included low and high copy number HCV armoured RNA used during extraction and RT-PCR (Asuragen). A negative, no template control was also included in each RT-PCR run.

2.3. Primers and probes

HCV Gt 1a, 1b, 2, and 3a primers and probes were designed using Geneious v.6.1.7 (Biomatters) and purchased from IDT or Life Technologies Inc. Conserved regions were identified using Core, E1, and NS5B sequences from the Los Alamos National Laboratory (LANL) database (Kuiken et al., 2005) and BCCDC (Olmstead et al., 2015). The total number screened were: six primers and three probes for Gt 1a, 12 primers and 13 probes for Gt 1b, 10 primers and six probes for Gt 3a, and two primers and one probe for Gt 2.

2.4. Real-time RT-PCR

Laboratory developed RT-PCR assays were performed on an ABI 7500 FAST enabled real-time PCR system with 1× Fast Virus 1-Step PCR Master Mix (Life Technologies Inc), 1× primer/probe mix (Table 1), and 5 µl nucleic acid. Thermocycling conditions were: 50 °C for 5 min, 95 °C for 20 s, then 40 cycles of 95 °C for 3 s and 60 °C for 30 s.



- Write in past tense
- Write in chronological order but group related methods together
- Usually have subheadings (great if match result subheadings)
- Don’t include any background or results

Results

Do: start with big picture description of methods

Do: use descriptive subheadings

Do: present data in past tense & logical order

Do: be clear and concise

Don't: put standard conditions in table

Don't: include everything, many tables/figures

"The fool collects facts; the wise man selects them"

- John Wesley Powell, President AAAS 1888

LPS modifications are not responsible for aminoglycoside resistance in *B. vietnamiensis*. To determine if LPS modifications are involved in aminoglycoside resistance in *B. vietnamiensis*, LPS compositions were compared between susceptible and resistant isolates. SDS-PAGE analysis of isolated LPS molecules revealed no gross differences among serial clinical isolates C8395, C8952, and D0774 from patient Bv1 and D0099 and D2075 from patient Bv2; all had rough LPS (LPS lacking O antigen) (data not shown). Overloading gels with up to 50 μg of LPS did not show the presence of O antigen in any of these isolates (data not shown). Lipid A species from the *B. vietnamiensis* isolates listed in Table 2 were analyzed by mass spectroscopy, and representative spectra are shown in Fig. 2A to F. Consistent with previous reports (25), lipid A structures were a blend of tetra- and penta-acylated molecules (Fig. 2A to F and data not shown). Lipid A structures of aminoglycoside-susceptible and -resistant *B. vietnamiensis* isolates were positive for Ara4N, identified on spectra by mass-to-charge ratios of 1,575, 1,601, 1,802, and 1,827 (Fig. 2A to F and data not shown).

Where are methods mentioned?

What data is omitted from presentation in full?

Results

Do: make sure your tables and figures can stand alone

Don't: repeat what is in the tables and figures

Don't: be verbose in citing tables and figures

It is clearly shown in Table 1 that the Check-Direct CPE and culture limit of detection (LoD) results were equivalent for seven (47%) isolates.

vs

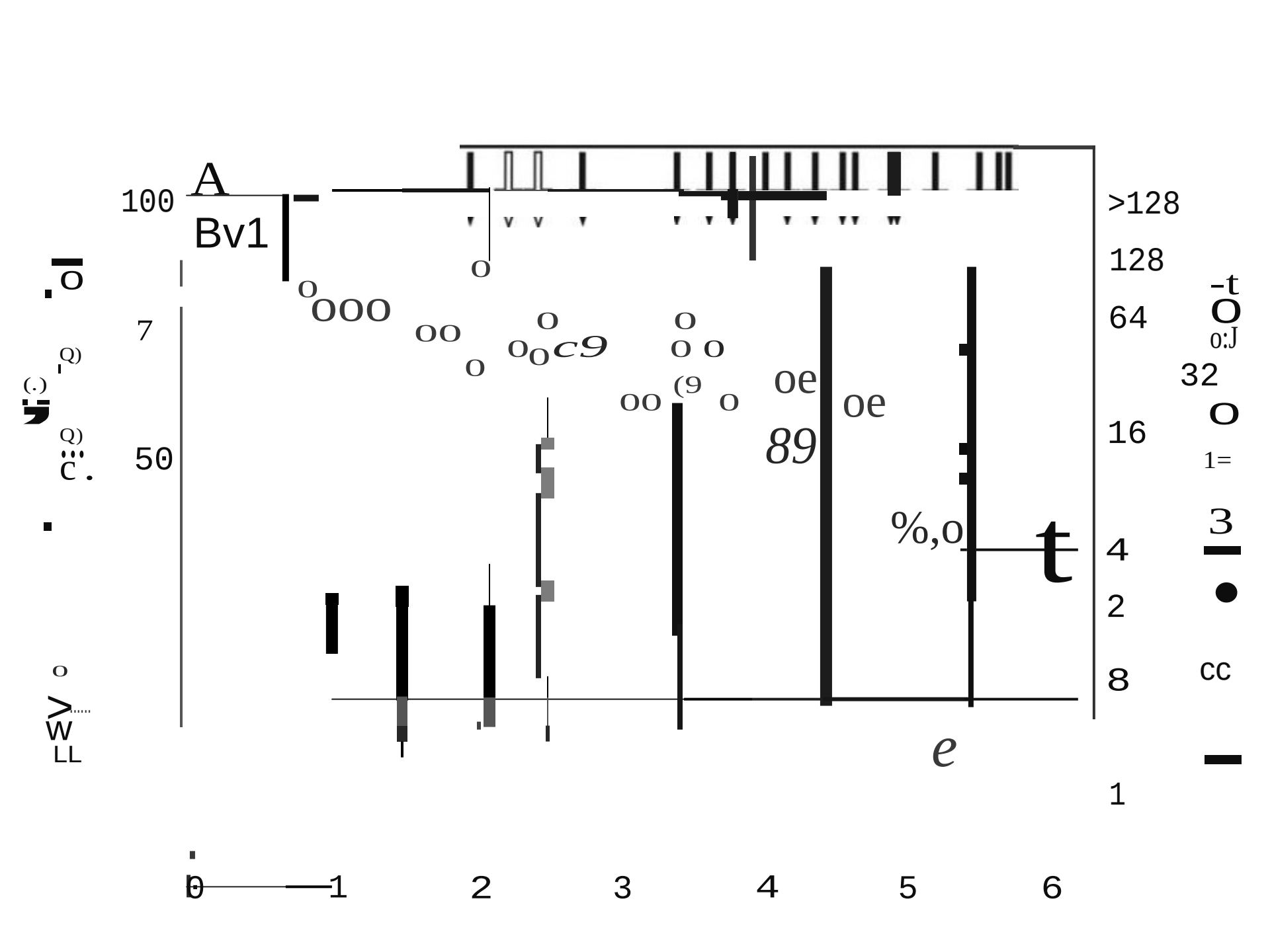
The Check-Direct CPE and culture limit of detection (LoD) results were equivalent for seven (47%) isolates (Table 1).

When to make a table

When you need to present the exact data but you can’t describe it easily in words

When to make a graph

When the data shows pronounced trends, making an interesting picture


A
Bv1
100
7
50
>128
128
64
32
16
4
2
8
1
0
1
2
3
4
5
6

Year

Formatting Tables

Like elements should read down

- **Do**: explain abbreviations
- **Do**: use only limited horizontal lines

Table 1
Fractional inhibitory concentration indices (FICIs) for combinations of novobiocin with other antibiotics against *Burkholderia cepacia* complex bacteria.

Species/strain	FICI for novobiocin +		
	PMB	Colistin	Tobramycin
B. cenocepacia K56-2	0.31	0.28	0.63
B. multivorans C0514	0.28	0.5	0.63
B. cenocepacia GIB-C1141144	0.5	0.56	>1.0
B. cenocepacia GRI-C1210932	0.31	0.38	>1.0

PMB, polymyxin B.

Discussion

Purpose: **show relationships among facts**

Be honest & clear & end with a bang!

1. Present a generalization of results
2. Summarize evidence for each conclusion
3. Point out exceptions or unsettled points
4. Show how your results agree or disagree with previously published work
5. Discuss theoretical implications and possible practical applications (don’t go too far)
6. State your conclusions

Overview

In this study, a paired, duplex real-time RT-PCR assay was developed and validated for identification of the most prevalent HCV genotypes in North America (1a, 1b, 2, and 3a). Even in the era of DAA, HCV genotype predicts treatment duration and efficacy (Gilead Sciences Inc, 2015). The RT-PCR assay was developed as an approach for low-cost, high-throughput detection of these major HCV genotypes to address an observed increase in genotype testing

Bonus! Present findings that didn't make the cut for the results section

Nakatani et al. (2010) identified subtypes 1a, 1b, 3a and 2a, 2b, 2c using two triplex RT-PCR tests, however, these primers and probes yielded false-positive results (Gt 4-positive samples reacted with 3a probe) when tested in the present study.

This qPCR screening approach may also be suitable for STEC detection in a diagnostic laboratory setting and surge capacity during a large-scale STEC outbreak, but additional quality requirements, such as inclusion of an internal control, must be explored. The potential benefits of the proposed pooled NAAT outweigh the observed analytical limitations, given that the method is a solution for overcoming current cost-based barriers associated with adoption of CDC recommendations for universal stool screening for non-O157 STEC. Public health interventions based on accurate estimates of STEC prevalence have the potential to reduce the burden of STEC infections.

Where are some limitations?

What is the significance?

Acknowledgements

Thanks to colleagues and funders

References

Include only relevant sources (this is not a review!)

TIPS for writing

- Don’t wait
- Find your zone
- Create a timeline with deadlines
- Manage your time

From Cirillo Company:
developed by Francesco Cirillo in the late 1980s

Work 25 mins, take 5 min break
At 4 pomodoros, take a longer break

TIPS for writing

- Focus
- Generate interest
- Tell a story
- Use examples
- Nail your contributions

Don’t: make readers do detective work

Don’t: write long-winded sentences “an adequate amount of = enough”

Don’t: use unnecessary words like “interestingly”

Do: use words like novel and innovative

TIPS for writing

Do: strategically repeat your main message in the Intro, Results, Discussion and Conclusion

Do: stick strictly to the main text word limit (~3000 words for a primary research paper), try to limit to 50 references

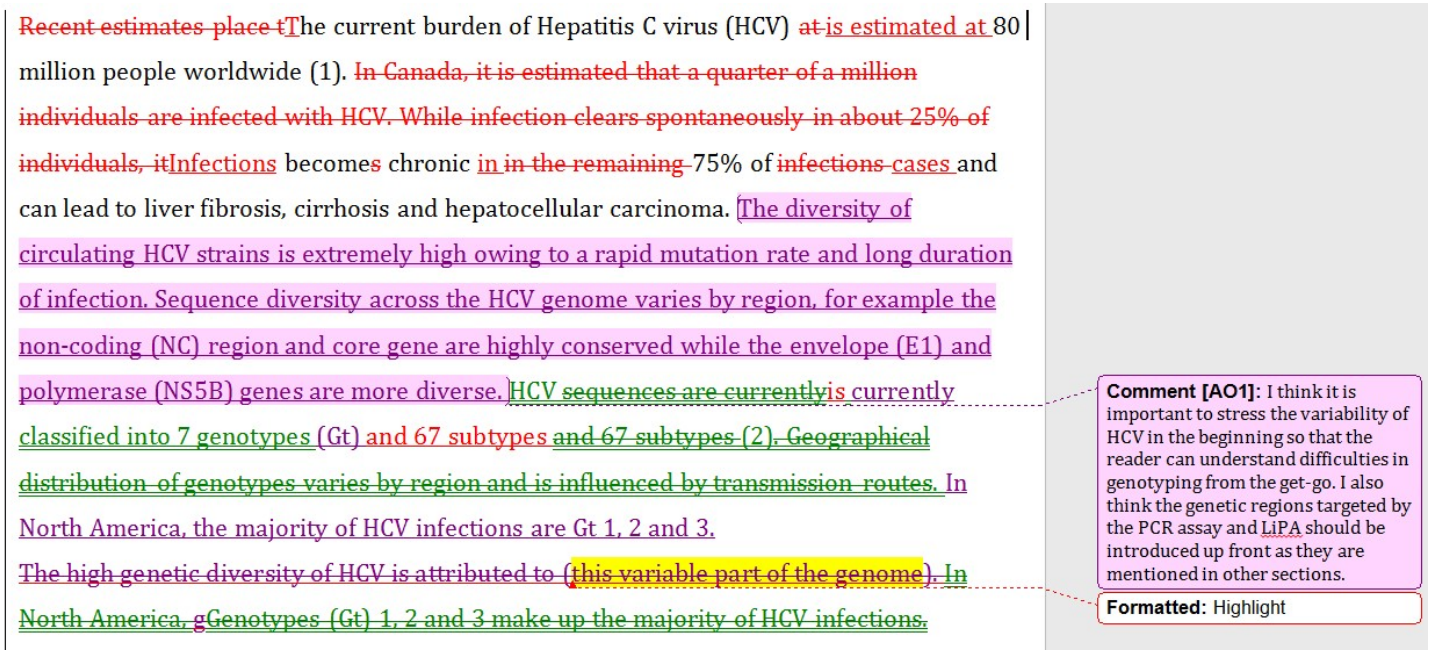
Don't: include more than 6 display items

TIPS for writing

Edit, Edit, Edit

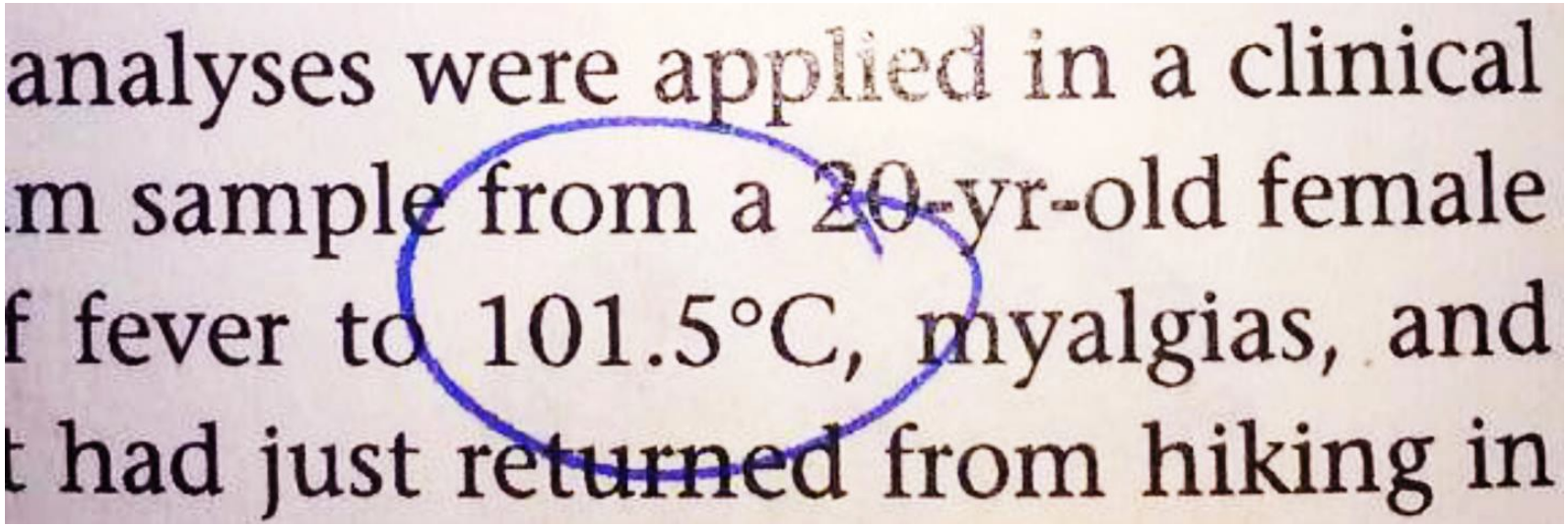
(don't wait, use colleagues)

Kill your darlings, kill your darlings, kill your darlings



TIPS for writing

Spell check



e.g. acid fast bacteria = acid-fast bacteria

Authorship

- Discuss early

4 ICMJE criteria:

- Substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work; AND
- Drafting the work or revising it critically for important intellectual content; AND
- Final approval of the version to be published; AND
- Agreement to be accountable for all aspects of the work in ensuring questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

16 digit number:

Permanent identifier for researchers

Consider if you have a common name or name change

Consider if not all publications found in one database

Can also track employment history etc

vworks (7)

Export works - Bulk edit

USort

LI KW ORKS

Hide lin(works

OROD works with our member organizations to make it easy to connect your ORCID iD and link to information in their records. Choose one of the link wizards to get started. More information about linking works

Work type Articles

Geograph ical area Global

CrossRef Metadata Search

Import your publications from crossRef's authoritative,publisher-supplied metadata on over 70 ...

Agatha jassem

E)orcid.org/0000-0002-4128-29 19

YWorks (7)

USort

Pooled Nucleic Acid Amplification Test for Screening of Stool Specimens for Shiga Toxin-Producing Escherichia coli

Journal of Clinical Microbiology

201698(journal-article

DOI: l0.1128/jcm.01373-16

Source: CrossRef Metadata Sea--ch

e Preferred source

Characterization ofCiostridium difficileStrains in British Columbia, Canada: A Shift from NAPl Majority (2008) to Novel Strain Types (2013) in One Region

Canadian Journal of Infectious Diseases and Medical Microbiology

2016 |jou rnal-article

DOI: l0.1 lSSn016/8207418

Source: CrossRef Metadata Sea--ch

e Preferred source

Submission: Cover Letter

Dear Journal of Clinical Microbiology Editorial Office,

Please find enclosed the manuscript entitled "Development and validation of a real-time, reverse transcription PCR assay for rapid and low-cost genotyping of hepatitis C virus genotypes 1a, 1b, 2, and 3a" for submission to the Journal of Clinical Microbiology.

The work outlined in this manuscript is of interest to clinical microbiology laboratories as it presents a low-cost and high-throughput strategy to address the burdens associated with an increased demand for Hepatitis C virus (HCV) genotyping. Selection of the optimal treatment regimen for HCV infection requires pre-determination of the infecting HCV genotype. The recent approval of direct acting antivirals for treatment of HCV has led to an increased demand for HCV genotyping. In the present study, we developed and validated an in-house reverse transcription PCR method for rapid and low-cost screening of the most prevalent HCV genotypes in North America.

No manuscripts related to this study have been published or submitted for publication elsewhere.

This manuscript has been seen and approved by all co-authors. All authors fulfill the authorship criteria. All authors report no potential conflicts of interest

Choosing a Journal

- Define your audience
- Match between with journal's aims and scope
- Aim high & for visibility
- Consider open access policy
- Consider publication turn-around-time
  - ASM mSphere
  - Eurosurveillance Rapid Communications

Choosing a Journal

Because of the large number of manuscripts submitted to the Journal of Clinical Microbiology that describe molecular methods, the editors decided that we would not encourage publication of manuscripts that essentially described the development of straightforward applications of recognized molecular techniques. It was determined that studies pertaining to molecular diagnostic methods needed to offer something that was truly novel in order to merit consideration for publication. Two editors for the journal have reviewed your paper. Given the extensive data on HCV genotyping methods, and the fact that laboratory-developed tests are not likely to be widely used, we felt that the novelty and impact of this work were limited.

Unfortunately, in view of this policy, we are not able to consider your paper for publication in JCM.

ubject Re: Suitibility of Manucript for Publication in JCV [161123-010480]

Dear Agatha,

Follov.ting your enquiry last week I have received a response from the Editor. On this occasion I'm afraid your paper isn't suitable for JCV, the Editor has advised it would be better sllited to a joumal of Virological methods.

Responding to reviewers

4	The authors may specify that these are catalog list prices.	104	In line # of the revised manuscript, we have specified that we are referring to catalog list prices.
5	the calculations are not clear as to what they are standardized against? What is the gold standard in which 100% was calculated against? Knowledge of at least one pool member being positive? This was not clear, if it was explained.	135	In the methods section of the revised manuscript, we explained the calculation in line 131: "The sensitivity values of pooling PCR strategies were calculated against individual PCR results". We also clarified how sensitivity was calculated in the Results section by changing this sentence at line 140 from "The overall sensitivity was 83% for the pre-enrichment

1) the limitation of using G-blocks (DNA) should be noted vs. viral RNA (not the natural target, but a suitable alternative). This should be discussed in the context of other possible options (in vitro transcription or high cost of custom armored RNA)

Response: The authors agree that there is a limitation to using DNA gBlocks as opposed to RNA options, but gBlocks are a cost efficient and convenient option for positive controls in limited analysis, such as precision. We have emphasized the limitation of gBlocks by adding the following sentence to lines 308-310 of the discussion: "Of note, the use of DNA gBlocks for positive controls and precision testing as opposed to RNA options such as HCV armored RNA is not ideal since HCV is an RNA virus, but gBlocks offer a cost effective and convenient alternative." Ultimately, the validity of the RT-PCR assay was confirmed using clinical samples with known viral loads and genotypes.

Revised Manuscript

Highlight revised portions mentioned in rebuttal:

308 and spiked samples. Of note, the use of DNA gBlocks for positive controls and precision
309 testing as opposed to RNA options such as HCV armored RNA is not ideal since HCV is an
310 RNA virus, but gBlocks offer a cost effective and convenient alternative.

Mention reviews in cover letter:

The reviews we received after our initial submission were very positive and we are happy to have the opportunity to re-submit a revised version of the manuscript that addresses reviewer concerns. No manuscripts related to this study have been published or submitted

Good luck
&
happy writing!

References

Professor Simon Peyton Jones, Microsoft Research
https://www.youtube.com/watch?v=g3dkRsTqdDA

How to Write & Publish A Scientific Paper, 5th ed
(given to me by my PI in my first year of gradschool)

conservationbytes.com/2012/10/22/how-to-write-a-scientific-paper/

My life (11 years doing research)