Factors Related to the Occurrence and Distribution of Select Bacterial and Protozoan Pathogens in Pennsylvania Streams

Joseph W. Duris
U.S. Geological Survey
Michigan Water Science Center
(517) 887-8942
jwduris@usgs.gov
Background

- What do indicators indicate?
- New studies dealing with fecal indicator bacteria (FIB) relation to pathogens
  - (Streams) Season, land use, discharge, turbidity, chemistry
  - (Lakes) Wave height, wind direction, storm drain flow
- Few studies have the same level of chemical and hydrologic data available from the WQN
Shiga toxin-producing *E. coli* (STEC)

- STEC are a main cause of intestinal disease in humans
  - Over 200 STEC types identified with various combinations of virulence genes
- STEC have been associated with outbreaks and death
  - Food and Water
- *E. coli* O157:H7 is most frequently isolated type in North America
  - Only STEC type that is “easy” to identify
  - Small infectious dose (fewer than 10 cells)
  - Typically most virulent combination of genes
- Microbial Source Tracking (MST) application
Pathogenic Enterococci

- Pathogenic *Enterococcus*
  - Not a common intestinal pathogen (i.e. does not typically cause diarrhea)
  - However, enters body via oral/fecal transmission
  - Enterococcus strains with the *esp* gene
    - Bacteremia
    - Endocarditis
    - Wound infection
    - Urinary Tract Infection
  - *esp* gene is cited in the literature as being a good marker for Enterococci from a *human* source (MST application)
Giardia and Cryptosporidium

- Analysis done by PA-DEP Bureau of Labs
- Protozoan parasites
- USEPA Method 1623
- **Giardia**
  - Causes Giardiasis, (diarrheal disease)
  - Exist as cysts in the environment (moderately chlorine resistant)
  - 6-19 microns
- **Cryptosporidium**
  - Causes Cryptosporidiosis (diarrheal disease)
  - Exist as oocysts in the environment (chlorine resistant)
  - 3-5 microns
Pathogen and MST Approach

- Use of fecal indicator bacteria cultures as platform
  - Growth/viability of the target organisms
    - Original Indicator Sample
  - Target pathogens and source tracking markers within those cultures
    - Enrichment PCR
    - Improved detection limits over direct PCR
      - Most direct PCR assays reported in the literature have high detection limits (e.g., 20 target CFU/ml) due to matrix interference and small volumes
  - Relevance to water quality criteria/standards
    - Pathogen occurrence can be related to concentrations of fecal indicator bacteria
## Pathogen and MST Gene Targets

<table>
<thead>
<tr>
<th>Organism</th>
<th>Gene Name</th>
<th>Gene Product</th>
<th>Gene Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>E. coli</td>
<td>eaeA</td>
<td>Intimin protein</td>
<td>Allows <em>E. coli</em> to tightly bind intestine</td>
</tr>
<tr>
<td>E. coli</td>
<td>stx1</td>
<td>Shiga Toxin 1</td>
<td>Causes less severe disease in humans (more common in ruminants than human)</td>
</tr>
<tr>
<td>E. coli</td>
<td>stx2</td>
<td>Shiga Toxin 2</td>
<td>Causes severe disease in humans (HC, HUS)</td>
</tr>
<tr>
<td>E. coli</td>
<td>E. coli 16s</td>
<td><em>E. coli</em> ribosome</td>
<td>Universal marker of all <em>E. coli</em>, control</td>
</tr>
<tr>
<td>E. coli</td>
<td>rfbO157</td>
<td>O157 surface protein</td>
<td>Marker of <em>E. coli</em> O157 (hamburger outbreak <em>E. coli</em>)</td>
</tr>
<tr>
<td>E. coli</td>
<td>LTIIa</td>
<td>Heat Labile Toxin Subunit II</td>
<td>Bovine Sources, gastroenteritis in humans</td>
</tr>
<tr>
<td>E. coli</td>
<td>STII</td>
<td>Heat Stabile Toxin Subunit II</td>
<td>Swine Sources, swine and human disease</td>
</tr>
<tr>
<td>Enterococci</td>
<td>esp</td>
<td>Enterococcus surface protein</td>
<td>Human Sources, associated with human disease (skin and bladder infections)</td>
</tr>
</tbody>
</table>
Measured Parameters

- Presence/absence of Gene Markers (Pathogen & MST)
- Densities of *E. coli*, enterococci, *Cryptosporidium* oocysts, *Giardia* cysts
- Physical parameters
  - Discharge, total dissolved solids, total suspended solids, ANC
- Chemistry
  - Basic chemistry
  - Nutrients
  - Metals
  - Organics
    - Pharmaceuticals, antibiotics, hormones
- Site Characteristics
  - Land-use, proximal & catchment, basin slope, drainage area, % carbonate bedrock and glacial deposit at catchment level
Study Area

- 27-stations
- Quarterly samples
- 2 years (2007-2009)
- Drainage areas from 5 km² to 49,000 km²
- Within 5 miles of drinking water intake
Study Goals

- Determine the frequency of occurrence of select bacterial and protozoan pathogens
- Determine relation of Indicator Bacteria to pathogens
- Determine factor(s) related to the occurrence of pathogens in Pennsylvania streams
Frequency of Pathogen Detection

![Bar chart showing the frequency of detection for different genes or pathogens](chart.png)

- eaeA
- stx2
- stx1
- EC
- rfb
- LTII
- STII
- esp
- Crypto
- Giardia

**Gene or Pathogen**

**Frequency of Detection (n=214)**
## Pathogen Relation to Indicator Criteria

<table>
<thead>
<tr>
<th></th>
<th>(n)</th>
<th>eaeA (%)</th>
<th>stx2 (%)</th>
<th>stx1 (%)</th>
<th>rfbO157 (%)</th>
<th>LTIIa (%)</th>
<th>STII (%)</th>
<th>esp (%)</th>
<th>Cryptosporidium (%)</th>
<th>Giardia (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Meet <em>E. coli</em> criteria</strong></td>
<td>186</td>
<td>32</td>
<td>7.6</td>
<td>3.2</td>
<td>3.8</td>
<td>3.8</td>
<td>1.6</td>
<td>2.2</td>
<td>44</td>
<td>58</td>
</tr>
<tr>
<td><strong>Exceed <em>E. coli</em> criteria</strong></td>
<td>31</td>
<td>97</td>
<td>53</td>
<td>26</td>
<td>32</td>
<td>32</td>
<td>23</td>
<td>42</td>
<td>61</td>
<td>77</td>
</tr>
<tr>
<td><strong>p-value</strong>&lt;sup&gt;b&lt;/sup&gt;</td>
<td>--</td>
<td>p&lt;0.05</td>
<td>p&lt;0.05</td>
<td>p&lt;0.05</td>
<td>p&lt;0.05</td>
<td>p&lt;0.05</td>
<td>p&lt;0.05</td>
<td>p&lt;0.05</td>
<td>p=0.100</td>
<td>p=0.057</td>
</tr>
<tr>
<td><strong>Meet enterococci criteria</strong>&lt;sup&gt;c&lt;/sup&gt;</td>
<td>163</td>
<td>29</td>
<td>5.5</td>
<td>2.5</td>
<td>3.1</td>
<td>1.2</td>
<td>0.6</td>
<td>1.8</td>
<td>42</td>
<td>57</td>
</tr>
<tr>
<td><strong>Exceed enterococci criteria</strong>&lt;sup&gt;c&lt;/sup&gt;</td>
<td>54</td>
<td>78</td>
<td>43</td>
<td>19</td>
<td>22</td>
<td>26</td>
<td>17</td>
<td>26</td>
<td>57</td>
<td>70</td>
</tr>
<tr>
<td><strong>p-value</strong>&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
<td>p&lt;0.05</td>
<td>p&lt;0.05</td>
<td>p&lt;0.05</td>
<td>p&lt;0.05</td>
<td>p&lt;0.05</td>
<td>p&lt;0.05</td>
<td>p&lt;0.05</td>
<td>p=0.076</td>
<td>p=0.115</td>
</tr>
</tbody>
</table>

<sup>a</sup> Recreational Water Quality Moderate Full Body Contact Criteria for *E. coli*, 298 CFU 100 L<sup>-1</sup>

<sup>b</sup> Pearson chi-square test with Yates’ correction

<sup>c</sup> Recreational Water Quality Moderate Full Body Contact Criteria for enterococci, 78 CFU 100 L<sup>-1</sup>

Preliminary Data, subject to revision
Seasonal Differences in Indicator Bacteria and Protozoan Densities and Frequencies

<table>
<thead>
<tr>
<th></th>
<th>Concentration</th>
<th>Frequency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n)</td>
<td>E. coli</td>
</tr>
<tr>
<td>Winter</td>
<td>26</td>
<td>295</td>
</tr>
<tr>
<td>Spring</td>
<td>84</td>
<td>296*</td>
</tr>
<tr>
<td>Summer</td>
<td>54</td>
<td>1186*</td>
</tr>
<tr>
<td>Autumn</td>
<td>50</td>
<td>435</td>
</tr>
</tbody>
</table>

p-value<sup>c</sup> -- p<0.05 p<0.05 P=0.270 p<0.05 <0.05 p=0.227 p=0.387 p<0.05

<sup>a</sup> [E. coli], Geometric Mean CFU100 mL<sup>-1</sup>; [Enterococci] Mean CFU100 L<sup>-1</sup>
<sup>b</sup> [Cryptosporidium] Mean oocysts, 1 L<sup>-1</sup>; [Giardia], Mean cysts, 1 L<sup>-1</sup>
<sup>c</sup>Pearson chi-square test with Yates’ correction (frequency), or ANOVA on ranks of log transformed concentrations (concentrations)

Note: Dec-Feb, Winter; Mar-May, Spring; Jun-Aug, Summer; Sep-Nov, Autumn

Preliminary Data, subject to revision
# Pathogen Relation to Season

<table>
<thead>
<tr>
<th>Season</th>
<th>(n)</th>
<th>eaeA (%)</th>
<th>stx2 (%)</th>
<th>stx1 (%)</th>
<th>rfbO157 (%)</th>
<th>lttlα (%)</th>
<th>stll (%)</th>
<th>esp (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Winter</td>
<td>26</td>
<td>38</td>
<td>15</td>
<td>3.8</td>
<td>3.8</td>
<td>7.7</td>
<td>7.7</td>
<td>15</td>
</tr>
<tr>
<td>Spring</td>
<td>84</td>
<td>24</td>
<td>9.5</td>
<td>4.8</td>
<td>7.1</td>
<td>6.0</td>
<td>2.4</td>
<td>7.1</td>
</tr>
<tr>
<td>Summer</td>
<td>54</td>
<td>59</td>
<td>23</td>
<td>12</td>
<td>11</td>
<td>11</td>
<td>7.1</td>
<td>8.9</td>
</tr>
<tr>
<td>Autumn</td>
<td>50</td>
<td>45</td>
<td>14</td>
<td>3.9</td>
<td>7.8</td>
<td>5.9</td>
<td>3.9</td>
<td>3.9</td>
</tr>
<tr>
<td><strong>p-value</strong>&lt;sup&gt;a&lt;/sup&gt;</td>
<td>--</td>
<td><strong>p&lt;0.05</strong></td>
<td>p=0.259</td>
<td>p=0.414</td>
<td>p=0.9127</td>
<td>p=0.874</td>
<td>p=0.815</td>
<td>p=0.625</td>
</tr>
</tbody>
</table>

<sup>a</sup>Pearson chi-square test with Yates’ correction

Note: Dec-Feb, Winter; Mar-May, Spring; Jun-Aug, Summer; Sep-Nov, Autumn

Preliminary Data, subject to revision
# Relation of Discharge to Indicator and Protozoan Densities

<table>
<thead>
<tr>
<th>Discharge Category&lt;sup&gt;a&lt;/sup&gt;</th>
<th>(n)</th>
<th>E. coli&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Enterococci&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Cryptosporidium&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Giardia&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>35</td>
<td>49*</td>
<td>32*</td>
<td>0.08</td>
<td>0.41</td>
</tr>
<tr>
<td>Median</td>
<td>118</td>
<td>333</td>
<td>389</td>
<td>0.13</td>
<td>1.27*</td>
</tr>
<tr>
<td>High</td>
<td>61</td>
<td>1242*</td>
<td>1441*</td>
<td>0.15</td>
<td>0.37*</td>
</tr>
<tr>
<td>p-value&lt;sup&gt;c&lt;/sup&gt;</td>
<td>--</td>
<td>p&lt;0.05</td>
<td>p&lt;0.05</td>
<td>p=0.224</td>
<td>p&lt;0.5</td>
</tr>
</tbody>
</table>

<sup>a</sup> Low, discharge measured was less than 25<sup>th</sup> percentile of daily mean stream flow at sample location; Median, discharge measured was between 25<sup>th</sup> and 75<sup>th</sup> percentile of daily mean stream flow at sample location; High, Discharge measured was greater than 75<sup>th</sup> percentile of daily mean stream flow at sample location.

<sup>b</sup> [E. coli], Mean CFU100 L<sup>-1</sup>; [Enterococci] Mean CFU100 L<sup>-1</sup>; [Cryptosporidium] Mean oocysts, 1 L<sup>-1</sup>; [Giardia], Mean cysts, 1 L<sup>-1</sup>

<sup>c</sup> Pearson chi-square test with Yates’ correction (frequency), or ANOVA on ranks of log transformed concentrations (concentrations)

---

Preliminary Data, subject to revision
## Relation of Discharge to Pathogen Frequencies

<table>
<thead>
<tr>
<th>Discharge Category&lt;sup&gt;a&lt;/sup&gt;</th>
<th>(n)</th>
<th>eaeA</th>
<th>stx2</th>
<th>stx1</th>
<th>rfbO157</th>
<th>ltIIa</th>
<th>stII</th>
<th>esp</th>
<th>Crytopsporidium</th>
<th>Giardia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>35</td>
<td>31</td>
<td>5.7</td>
<td>2.9</td>
<td>2.9</td>
<td>5.7</td>
<td>0</td>
<td>0</td>
<td>34</td>
<td>60</td>
</tr>
<tr>
<td>Median</td>
<td>118</td>
<td>39</td>
<td>10</td>
<td>5.1</td>
<td>5.1</td>
<td>5.9</td>
<td>2.5</td>
<td>6.8</td>
<td>47</td>
<td>54</td>
</tr>
<tr>
<td>High</td>
<td>61</td>
<td>51</td>
<td>30</td>
<td>12</td>
<td>16</td>
<td>11</td>
<td>11</td>
<td>15</td>
<td>52</td>
<td>75</td>
</tr>
<tr>
<td>p-value&lt;sup&gt;c&lt;/sup&gt;</td>
<td>--</td>
<td>p=0.139</td>
<td>p&lt;0.05</td>
<td>p=0.164</td>
<td>p&lt;0.05</td>
<td>p=0.580</td>
<td>p&lt;0.05</td>
<td>p&lt;0.05</td>
<td>p=0.222</td>
<td>p&lt;0.05</td>
</tr>
</tbody>
</table>

<sup>a</sup> Low, discharge measured was less than 25th percentile of daily mean stream flow at sample location; Median, discharge measured was between 25th and 75th percentile of daily mean stream flow at sample location; High, Discharge measured was greater than 75th percentile of daily mean stream flow at sample location

<sup>b</sup> [E. coli], Mean CFU100 L<sup>-1</sup>; [Enterococci] Mean CFU100 L<sup>-1</sup>; [Cryptosporidium] Mean oocysts, 1 L<sup>-1</sup>; [Giardia], Mean cysts, 1 L<sup>-1</sup>

<sup>c</sup> Pearson chi-square test with Yates’ correction (frequency), or ANOVA on ranks of log transformed concentrations (concentrations)

Preliminary Data, subject to revision
Pathogen Relation to Likely Source

† Significant difference p<0.05

Preliminary Data, subject to revision
“Data Exploration”

- Original data set had 214 samples and over 220 measured parameters and sampling location observations
- Eliminated measurements that were mostly non-detects, or were incomplete across all sampling years
  - Some substitution of censored data according to established procedures.
- Cluster analysis and principle components analysis with remaining observations
Cluster Analysis

Low Pathogen cluster
- Less frequent pathogen gene detection
- Greater proximal forest, and less proximal urban land use
- Less carbonate bedrock area
- Fewer indicator bacteria
- Lower suspended solids
- Lower total nitrogen
- Lower caffeine concentrations
Cluster Analysis

Intermediate Pathogen cluster

- Significantly greater upstream agricultural land use, and area of carbonate bedrock
- Significantly different (intermediate) indicator bacteria and suspended solids
- Significantly greater total nitrogen, carbamazepine, and sulfamethoxazole concentrations

Preliminary Data, subject to revision
Cluster Analysis

High Pathogen Cluster
- Mixed land use
- Significantly different (intermediate) carbonate bedrock
- Significantly greater drainage area and unit discharge
- Significantly greater indicator bacteria and suspended solids
- Greater Fe, Mn, Al, and acetaminophen concentrations

Preliminary Data, subject to revision
Conclusions

- Bacterial and protozoan pathogens were frequently detected.
- Samples exceeding standards are more likely to contain bacterial pathogens (not protozoa).
- Certain pathogenic populations of EC and ENT are affected by increased discharge while other populations of EC, most likely driven by source, are unaffected by increased discharge.

Preliminary Data, subject to revision
Conclusions

- The presence of MST markers was a poor discriminator of EC pathogens.
- *Giardia* frequency increased in winter and spring under median flow conditions.
  - *Giardia* occurrence is related more to non-point sources that are highly influential during seasonal overland transport resulting from snow melt and spring rain.
- More frequent pathogen detection was associated with increased concentrations of acetaminophen, AI, Fe, and Mn.
  - Suggests more frequent pathogen occurrence related to human waste and practices.
Thank You.

- PA-DEP
- PA-DEP BOL
- USGS PA-WSC
- Co-authors:
  - Andrew G. Reif, USGS PA-WSC
  - Donna A. Krouse, PA-DEP BOL
  - Natasha Cosgrove, USGS MI-WSC

*Factors Related to the Occurrence and Distribution of Select Bacterial and Protozoan Pathogens in Pennsylvania Streams. In revision for submission to Water Research

Preliminary Data, subject to revision
Factors Related to the Occurrence and Distribution of Select Bacterial and Protozoan Pathogens in Pennsylvania Streams

Joseph W. Duris
U.S. Geological Survey
Michigan Water Science Center
(517) 887-8942
jwduris@usgs.gov