ISSUES IN SETTING SECONDARY CONTACT RECREATION GUIDELINES 
FOR NEW ZEALAND FRESHWATERS

Graham McBride, NIWA, Hamilton, NZ (Graham.McBride@niwa.co.nz)
9 September 2012

Summary

The freshwater component of the MfE/MoH Microbiological Water Quality Guidelines for Marine and Freshwater Recreational Areas (MfE/MoH 2003) is based on a Quantitative Microbial Risk Assessment (QMRA) model for campylobacteriosis infection among swimmers. This QMRA model has been reconfigured with reduced levels of exposure, to mimic secondary contact recreation (canoeing, boating and fishing). Results for five different levels of exposure reduction are presented, to better enable informed choice in the light of new information about levels of exposure during secondary contact recreation. These all result in increased E. coli breakpoints. Included is a new E. coli breakpoint corresponding to a 0.5% tolerable infection risk, in addition to the current Guidelines’ “tolerable” levels of infection risk: 0.1%, 1% and 5% in the boundaries between its “Microbiological Assessment Categories”: A-B, B-C, C-D. These breakpoints could be used to define the boundaries between “excellent”, “good”, “fair” and “poor” bands. The manner in which site E. coli data can be used to assess which band a site should occupy raises some important and unresolved questions, particularly with respect to accounting for variability of bacteria concentrations (E. coli and Campylobacter) at a site, and attainability.

Fundamental approach to recreational water management

Two issues deserve consideration.

(1) The MfE/MoH Guidelines (2003) are built on twin objectives of a monitoring programme for recreational areas: Surveillance and Grading. These invoke short-term and long-term considerations, respectively. The former refers to issues such as occur when the most recent data indicate the possibility of a problem needing attention (e.g., unexpected wastewater overflows impinging on a swimming area). The latter refers to the general conditions obtaining over time at a recreation area, and includes the notion of reviewing the status of that area as the years go by. This duality seems to be quite harmonious with the objectives of the Land and Water Forum. So I suggest that it has much to commend itself to National Objectives Framework activities.

(2) The burden-of-proof. The compliance assessment rule for the New Zealand Drinking-Water Standards was derived using a precautionary approach to the burden of proof. The same precaution has been used as has been used when setting the E. coli breakpoints for the primary contact recreation in the freshwater component of the Guidelines. The coastal waters component of the Guidelines is based on epidemiological studies carried out in the UK using enterococci as an illness risk indicator, as explained by Kay et al. (2004). Applying the same Guidelines to freshwater [as is possible under the WHO (2003) Guidelines], was not considered appropriate for New Zealand: (a) because enterococci had been found in high concentrations in catchment headwaters in this country, and (b) because most of the (few) freshwater epidemiological studies that had been done were for lakes, not rivers.

Note that the Guidelines use both “acceptable” and “tolerable” when referring to health risk. The latter seems the more appropriate—and is particularly endorsed at section A.1 (“Health Risks”) at page A2.
MfE/MoH Guidelines. Whether a similar approach should be taken with a compliance assessment rule for secondary contact recreation is an open question (e.g., expressed as values of sample a medians or sample 95%iles).³

Consequences of breaching compliance

The recreational water guidelines deliberately avoid a punitive approach. Instead, if surveillance data indicate a problem then adaptive management actions are required, especially: increased sampling, public advisories, and catchment sanitary surveys. Also, if successive long-term Grading indicate deterioration, its causes then require attention.

Derivation of existing risk profiles for primary contact recreation

The Freshwater Microbiological Guidelines (MfE/MoH 2003, Table H2, page H26) contain 95%ile E. coli breakpoints for beach grading for primary contact recreation. These were obtained using a QMRA (Quantitative Microbial Risk Assessment) model for campylobacteriosis.⁴ ⁵ QMRA takes explicit account of variability in key inputs, particularly exposure and pathogen concentration, to build a risk profile (cf. a single risk value), using Monte Carlo statistical modelling techniques (Haas et al. 1999, McBride 2005). The QMRA underpinning the Guidelines used the @RISK Excel plug-in Monte Carlo software (Palisade Corp., 2000).

The model uses Campylobacter stream and lake data from the Freshwater Microbiological Research Programme Report (FMRPR, McBride et al. 2002, Till et al. 2008) and the campylobacteriosis infection dose-response curve for adults (Medema et al. 1996)—based on clinical trial data reported by Black et al. (1988). The FMRP measured E. coli and Campylobacter concentrations (and other pathogens and indicators) at 22 river sites and 3 lake sites spread over New Zealand. Sampling was performed fortnightly for 15 months in 1998-2000.

How much exposure should we assume for secondary contact?

This is the key issue for this analysis, because the only the exposure variables have been changed (i.e., reduced) when using the FRMP campylobacteriosis QMRA model to assess health risks associated with secondary contact. That recognises the distinguishing feature of secondary contact recreation: reduced exposure.

³ After all, once samples are taken one should consider this: What summary statistic of the data collected should be compared to the numerical objective? The expression of the numerical objective should make that clear. Other systems (e.g., the ANZECC Guidelines) don’t.
⁴ QMRA is by now a well-established technique—Haas et al. (1999), Rijal (2010), McBride et al. (2012).
⁵ Campylobacteriosis was selected for QMRA analysis because: (i) it has New Zealand’s highest reported rate (currently about 200 per 100,000 people per annum), (ii) it can be waterborne (as well as foodborne)—Till & McBride (2004); (iii) it has commonly been found in New Zealand recreational freshwaters (Till et al. 2008, French et al. 2011); and (iv) its dose-response characteristics are reasonably well understood. And now we can add the finding that ovine and bovine Campylobacter strains seem now to be dominant (over poultry strains) in human stools (French et al. 2011). Its details were guided by the “Freshwater Guidelines Advisory Group”, set up by MfE and MoH.
First, let us consider what exposure assumptions were used for primary contact recreation in the QMRA model. These are shown in Table 1.

Table 1: Exposure variables and statistics assumed for primary contact recreation

<table>
<thead>
<tr>
<th>Variable</th>
<th>Statistic (for assumed triangular distributions)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Minimum</td>
</tr>
<tr>
<td>Ingestion rate (mL/h)</td>
<td>10</td>
</tr>
<tr>
<td>Duration (h)</td>
<td>0.25</td>
</tr>
</tbody>
</table>

*Source: McBride et al. (2002), Till et al. (2008).*

The origin of the ingestion rates is explained in section A.3.7 of McBride et al. (2002, page 77, footnote 79): “The upper limits refer to swimmers; other recreational users (water skiers, wind surfers) tend to ingest or inhale less water (G. Lewis, University of Auckland, pers. comm.). Note too that Schernewski & Jülich (2001) have noted that “10 ml to 100 ml water are incorporated during bathing (Johl et al. 1995).”"

The duration figures (in Table 1) represent “Best Professional Judgement”.

Since that time (2002), and very recently, we have become aware of some further existing information and some that is more recent.

- Fewtrell et al. (1992) compared health effects between two groups of canoeists, one using lowland river water, the other using pristine upland waters. They found strong differences in self-assessed illness rates between the two groups, implying that ingestion of contaminated water was the likely cause.
- Fewtrell et al. (1994) report that approximately 8% of canoeists at freshwater sites reported capsizing and approximately 16% of rowers reported water ingestion. These studies indicated that these activities are likely to involve some degree of incidental water ingestion, some of which may not have been noticed.
- Rijal et al. (2011) cited two reports (Genthe & Rodda 1999, Medema et al. 2001) saying that: “A value of 10 mL/event was reported for the accidental gulping of water during activities such as cleaning laundry, fishing and agricultural/horticultural irrigation”.
- Schijven & de Roda Husman (2006) report that occupational divers, on average, swallowed 9.8 mL marine water and 5.7 mL fresh surface water per dive. Sport divers swallowed, on average, 9.0 mL marine water; 13 mL fresh recreational water; 3.2 mL river, canal, or city canal water; and 20 mL water in circulation pools. Some reported maximum ingestion of 190 mL per dive. These data were derived from questionnaires filled out by divers, and so are rather subjective.
- Dufour et al. (2006) report results of a pilot study in which ingestion of swimming pool water was estimated by the amount of cyanuric acid in a 24 hour urine sample collected after a one hour swimming activity in a pool disinfected with chlorine isocyanurate. The average amount of water swallowed by the 53 participants was about 30 mL. Swimmers less than 16 years old swallowed about 37 mL of water, which was more than twice that swallowed by adults (average 16 mL). Similar results were obtained in the follow-up study (Evans et al. 2006).
Rijal et al. (2011) attempted to quantify secondary contact ingestion rates for the Chicago Area Waterways System project. They present the data shown in Table 2 (copied from their Table 3).

Dorevitch et al. (2011) performed two studies in the Chicago area in 2009, one for self-assessed water ingestion and another using the cyanuric acid approach of Dufour et al. (2006). They found that the “Mean and upper confidence estimates of water ingestion during limited-contact recreation on surface waters are about 3–4 mL and 10–15 mL, respectively”.

Table 2: Exposure variables assumed for the Chicago Area Waterways System.

<table>
<thead>
<tr>
<th>Percentile</th>
<th>Boating (mL/h)</th>
<th>Fishing (mL/h)</th>
<th>Canoeing (mL/h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>1.49</td>
<td>2.98</td>
<td>5.21</td>
</tr>
<tr>
<td>25</td>
<td>1.65</td>
<td>3.30</td>
<td>6.02</td>
</tr>
<tr>
<td>50</td>
<td>1.90</td>
<td>3.79</td>
<td>7.52</td>
</tr>
<tr>
<td>75</td>
<td>2.23</td>
<td>4.47</td>
<td>10.15</td>
</tr>
<tr>
<td>90</td>
<td>2.64</td>
<td>5.28</td>
<td>14.16</td>
</tr>
<tr>
<td>95</td>
<td>2.95</td>
<td>5.89</td>
<td>17.84</td>
</tr>
<tr>
<td>97.5</td>
<td>6.26</td>
<td>6.51</td>
<td>21.99</td>
</tr>
<tr>
<td>100</td>
<td>7.43</td>
<td>22.13</td>
<td>34.00</td>
</tr>
</tbody>
</table>

Source: Rijal et al. (2011). These data were generated by assuming lognormal distributions of ingestion rates with assumed means and standard deviations (e.g., for canoeists these statistics were both assumed to be 5 ml/h).

To mimic secondary contact recreation I have run the primary contact QMRA model (using the latest @RISK software, Palisade Corp., 2009) with a series of reduced exposures set to: one half, one third, one quarter, one fifth and one tenth of the primary contact values. Results are shown on Table 3, along with E. coli percentiles for the FMRPR.

It remains to decide on the appropriate level of exposure reduction to be assumed for secondary contact, in the light of the above literature on exposure rates. (I suggest leaving the duration of exposure statistics unchanged.) Comparing the results reported by Rijal et al. (2011) and by Dorevitch et al. (2011) could suggest taking secondary exposure ingestion rates as one tenth of those used in the FMRPR for primary contact recreation (given in Table 1).
Table 3. Calculated risk profiles for primary and secondary contact recreation

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Water contact category (infection cases out of 1000)</th>
<th>E. coli / 100 mL (FMRPR)*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Primary (1/2)</td>
<td>Secondary (1/3)</td>
</tr>
<tr>
<td>5%ile</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10%ile</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>15%ile</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>20%ile</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>25%ile</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>30%ile</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>35%ile</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>40%ile</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>45%ile</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>50%ile</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>55%ile</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>60%ile</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>65%ile</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>70%ile</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>75%ile</td>
<td>18</td>
<td>7</td>
</tr>
<tr>
<td>80%ile</td>
<td>26</td>
<td>13</td>
</tr>
<tr>
<td>85%ile</td>
<td>72</td>
<td>40</td>
</tr>
<tr>
<td>90%ile</td>
<td>131</td>
<td>88</td>
</tr>
<tr>
<td>95%ile</td>
<td>329</td>
<td>330</td>
</tr>
</tbody>
</table>

* From the first column of results in Table A3.7.3 in McBride et al. (2002)

Interpretation of Table 3

The risk numbers in the cells are the number out of 1000 people who may become infected following ingestion of contaminated water at a swimming site on any particular day. The percentile column refers to percentiles of time. So for example, for primary contact recreation, 75% of the time no more than 18 out of 1000 exposed people would become infected.

Calculating E. coli breakpoints

The primary contact E. coli breakpoints were calculated (by the FMRP analyst) from the agreed tolerable risk levels at the three boundaries between the four Microbiological Assessment Categories (Table E2 of MfE/MoH 2003): A-B; B-C, C-D. Those breakpoints, as given in Table H2 (and justified on page I17), are 0.1%, 1% and 5%. So for primary contact recreation, the risk got to 0.1% as soon as 1 case of infection (out of 1000) was encountered. That happened between the 55%ile and 60%ile of the risk profile. The corresponding 55%ile for E. coli (which, at higher concentrations, is moderately well

Note that this value is higher than the primary value. This is just the result of the "slings and arrows of outrageous fortune". That is, counter-intuitive results can occur when conducting random sampling.
correlated with *Campylobacter*) is 131 per 100 mL. This was rounded down to 130 per 100 mL. Similar reasoning applies to the other risk levels. Note that in the case of the 5% risk level, the infection risk rose above 5% about halfway between the 80%ile and 85%ile. So, for that risk level, the approximate average of the 80%ile and 85%ile *E. coli* concentration was taken as the breakpoint *E. coli* value (i.e., 550 per 100 mL).

Note that the first meeting of the National Objectives Framework Science Panels (2 August, 2012) called for a further tolerable risk boundary, between 0.1% and 1.0%. Taking that value as 0.5%, the primary contact risk breakpoint would be about the 75%ile, in which case the risk breakpoint would be at 220 per 100 mL.

Table 4 presents these (approximate) breakpoints for the four new secondary contact categories as well.

<table>
<thead>
<tr>
<th>Breakpoints for tolerable infection risk</th>
<th>Primary <em>E. coli</em> breakpoints (E. coli per 100 mL)</th>
<th>Secondary <em>E. coli</em> breakpoints (E. coli per 100 mL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1%</td>
<td>130</td>
<td>190</td>
</tr>
<tr>
<td>0.5%</td>
<td>220</td>
<td>290</td>
</tr>
<tr>
<td>1%</td>
<td>260</td>
<td>395</td>
</tr>
<tr>
<td>5%</td>
<td>550</td>
<td>650</td>
</tr>
</tbody>
</table>

NB: Breakpoints are rounded to some extent, as seems appropriate for their intended use.

Note that reductions in exposure do not cause a concomitant linear increase in predicted *E. coli* breakpoints. For example, halving the exposure causes less than a doubling of the *E. coli* breakpoint, etc. In fact this concomitant increase does occur for the *Campylobacter* concentrations (as it should), but this simple relationship is confounded by the nonlinear relationship between *Campylobacter* and *E. coli* concentrations. In particular, in the FMRP data the distribution of *Campylobacter* concentration is more strongly right-skewed than the equivalent distribution of *E. coli* so that, for example, decreasing the primary exposure by a factor of four leads to less than four times the *E. coli* breakpoint.

**What about children?**

Available evidence, particularly from a large New Zealand campylobacteriosis modelling study (McBride *et al.* 2011) shows that children’s reported rate of campylobacteriosis is higher than for adults. It seems clear that, in part at least, this is because they are more exposed (i.e., ingest at a higher rate, and for a longer duration). That aspect is captured in the QMRA models, because a range of exposures is used. It may also be the case that children are more susceptible, i.e., exhibit a different dose-response curve. For example, an

---

7 If an individual’s exposure is halved and the ingested *Campylobacter* concentration is doubled (for a secondary exposure case), the dose would be the same as for a primary contact individual ingesting the actual *Campylobacter* concentration.

8 In part this can reflect the higher probability that parents will take an ill child to the doctor, whereas an adult might choose to “ride out the storm” when similarly afflicted. However, the New Zealand data also show that rural children report higher rates than urban children (Lake *et al.* 2011) and there is no reason to expect any strong difference in doctor referral rates between these two groups.
analysis of two campylobacteriosis outbreaks among primary school children (8–13 years old in The Netherlands, and 3–4 years old in England) who consumed milk apparently contaminated with *Campylobacter* showed that their ID$_{50}$ for illness was much lower than the ID$_{50}$ for infection among adults displayed in Figure 1 (Teunis *et al.* 2005). This finding needs to be tempered by the fact that no *Campylobacter* assays were carried out for the consumed milk and so the analysis is rather reliant on assumptions about that. Also, the medium of exposure (milk, cf. water) may have had an influence.

**Uncertainties**

The major uncertainties in the QMRA models include:

- *Dose response at low doses.* As shown on Figure 1, the lowest dose administered in the clinical trial was 800 *Campylobacter*;
- *Children’s increased susceptibility* (as discussed above);
- *Infectiousness of the Campylobacter species reported in the FMRPR.* The QMRA underpinning the 2003 guidelines assumed that all the *Campylobacter* measured in streams and lakes were equally infectious. Recent results (French *et al.* 2011) indicate that this may not be so, particularly for avian strains that can occur in stream/river water (particularly Pukeko).
- *The possible presence of other pathogens.* This is recognised in the guidelines (MoH/MfE 2003, at page I17), where the three (gastrointestinal illness) risk breakpoints recommended by WHO (2003) for marine waters (1%, 5%, 10%) have been lowered to 0.1%, 1% and 5%.

Note an apparent difficulty that may be dismissed. That is, the FMRP data for *Campylobacter* concentrations used in the QMRA models were for samples taken from recreational sites. Waters for which a secondary contact recreation guideline might apply may be more contaminated. However, that only implies that it may be more difficult than otherwise imagined for these waters to comply with secondary contact recreation guidelines. It doesn’t compromise the QMRA.

**What statistic(s) should be used to express the secondary contact breakpoints?**

This needs careful thought!

The primary contact guideline breakpoints (in MfE/MoH 2003) in Table 2 are “best guesses”. So it may be thought that an “even-handed” approach to their implementation would be to express these numbers as medians. But they are expressed as 95%iles (in keeping with the general approach recommended by WHO 2003). That is, the “even-handed” approach was deliberately not used.

That was because a precautionary approach was adopted to account for “statistical sampling error”, as follows. Consider the “best estimate” of the *E. coli* cut-off value for a 1% risk of campylobacteriosis infection for primary contact: 260 *E. coli* per 100 mL. Now consider using a median of values from a site’s sampling record as a “best estimate” for the true (population) median concentration of *E. coli* for that site, and require it to be below 260 per 100 mL in order to infer that the campylobacteriosis risk is less than 1%. If the true (but
Figure 1. Campylobacteriosis infection curve for adult volunteers, as derived by Medema et al. (1996)\(^9\) using clinical trial data of Black et al. (1988), in which the lowest average dose given to any group was 800 (as shown by the dot).

unknown) median *E. coli* concentration were to have been >260 per 100 mL, then the probability of getting a median of site data to be less than 260 per 100 mL could be as high as 0.5 (especially in borderline cases where the true *E. coli* concentration is only a little above 260 per 100 mL). Therefore, a precautionary approach was taken. Using 95\%iles, the probability of getting sample 95\%ile < 260 when the true median was > 260 is very low (McBride & Ellis 2001; McBride 2003).

However, my colleague Dr Sandy Elliott has observed that the choice of sample statistic(s) to compare with the suggested breakpoints (and so decide whether the site should be classed as “Excellent”, “Good”, “Fair” or “Poor”) may call for a more detailed examination of the variability of *Campylobacter* (or its surrogate, *E. coli*) at sites where data is available. As a result of that analysis it may even be possible to calculate risks directly, or to provide a simple lookup table the key sample statistic (or statistics) with which to compare against the breakpoint value.

Finally, note that the 2003 Guidelines confine sampling to “recreational areas” and to the bathing season, generally with weekly sampling, i.e., about 20 samples per year. Assessment of the Guidelines’ “Suitability for Recreation Grade” for those areas calls for assessment against 95\%ile values using five years of data (and for reassessments once more than five years’ data are to hand). However, many Regional Councils adopt a monthly

\(^9\) Figure 1 is a “beta-Poisson” curve, given by \(\text{Prob}_{\text{infection}} = 1 - (1 + d \beta)^{-\alpha}\), where \(d\) represents the average doses given to each group in the trial, \(\alpha (= 0.145)\) is a shape parameter and \(\beta (= 7.58)\) is a scale parameter. Note that this functional form is not arbitrary; it is derived from fundamental assumptions (such as the “single-hit hypothesis”) concerning the likelihood of at least one ingested *Campylobacter* surviving the body’s defences to cause infection. (For campylobacteriosis the beta-Poisson curve is a good approximation to the difficult-to-handle exact result—a Kummer hypergeometric function—it is not a good approximation for more infectious pathogens.)
sampling frequency for the water quality sites outside of designated primary recreational areas. If annual assessments of secondary water contact recreation limits are to be used (serving the role of general microbiological water quality condition) using 95%iles is fraught with danger.\textsuperscript{10} So if 95%iles are to be used for secondary contact recreation guidelines, it would be wise to define multi-year compliance assessment periods.

Attainability?

Figure 2 displays maps of predicted present-day median \textit{E. coli} concentrations and 95%ile \textit{E. coli} concentrations. These predictions are essentially based on all the relevant \textit{E. coli} records to data (at sites marked by dots) extrapolated using a random forests regression model. These may need consideration when deciding on a compliance rule for secondary contact.

The 95%ile map (right-hand panel in Figure 2) shows that a large proportion the streams in the country would receive a ‘poor’ rating in relation to primary recreation (red and yellow areas, plus part of the green in the right-hand diagram), and many areas would also be ‘poor’ in relation to the “1/4” secondary breakpoint (yellow and red areas). Hence the secondary breakpoints, and even the primary ones, are likely to be difficult to attain across the country. The 50%ile map (left-hand panel) shows that the same breakpoint concentration values would be more attainable, highlighting the sensitivity to the choice of percentile value. It may be useful in the future to have maps with the colours corresponding to proposed breakpoints, as is possible (pers. comm. Martin Unwin, NIWA, Christchurch).

\textsuperscript{10} 95%iles can be estimated using 12 data, if one uses the “Hazen” percentile estimator (see \url{http://search.mfe.govt.nz/search?q=hazen&site=MfE&client=MfE&proxystylesheet=MfE&output=xml_no_dtd}. But the resulting estimate will have considerable uncertainty.
Figure 2. Predicted present-day year-round *E. coli* concentrations using the random forests regression modelling method.\textsuperscript{11}

\textsuperscript{11} The left panel displays median values (Figure A2.4b in Unwin et al. 2010); the right panel displays 95\%iles (pers. comm. Dr Martin Unwin, NIWA, Christchurch). The floor and ceiling of each colour band shown on these figures have been selected automatically so as to have an adequate balance of colour in the maps; they bear no direct relationship to breakpoints being considered in this memo, and they are different for the 95\%ile and median values. The different density of black dots between the Figure’s panels indicates that more data were available to calculate 95\%iles.
REFERENCES


12 pdf file available from the author of this memo.


Palisade Corporation (2009). @RISK. Advanced Risk Analysis for Spreadsheets, v.5.5.0. Newfield, New York.


15 Not seen: cited by Rijal et al. (2011).