



---

James Mulhern, *President & Chief Executive Officer* | Randy Mooney, *Chairman*

July 29, 2015

Agri-Mark, Inc.  
Associated Milk Producers Inc.  
Bongards' Creameries  
Cooperative Milk Producers Association  
Cortland Bulk Milk Producers Cooperative  
Dairy Farmers of America, Inc.  
Dairymen's Marketing Cooperative, Inc.  
Ellsworth Cooperative Creamery  
Farmers Cooperative Creamery  
FarmFirst Dairy Cooperative  
First District Association  
Foremost Farms USA  
Land O'Lakes, Inc.  
Lone Star Milk Producers  
Maryland & Virginia Milk Producers Cooperative Association  
Michigan Milk Producers Association  
Mid-West Dairymen's Company  
Mount Joy Farmers Cooperative Association  
Northwest Dairy Association  
Oneida-Madison Milk Producers Cooperative Association  
Prairie Farms Dairy, Inc.  
Premier Milk Inc.  
Scioto County Cooperative Milk Producers' Association  
Select Milk Producers, Inc.  
Southeast Milk, Inc.  
St. Albans Cooperative Creamery, Inc.  
Swiss Valley Farms Company  
Tillamook County Creamery Association  
United Dairymen of Arizona  
Upstate Niagara Cooperative, Inc.  
Zia Milk Producers, Inc.

Division of Dockets Management, HFA-305  
Food and Drug Administration  
5630 Fishers Lane, Room 1061  
Rockville, MD 20852

**Re: Comments on Multicriteria-Based Ranking Model for Risk Management of Animal Drug Residues in Milk and Milk Products; Docket Number: FDA-2015-N-1305**

Dear Sir or Madam:

The National Milk Producers Federation (NMPF) thanks FDA for fulfilling the request from the National Conference on Interstate Milk Shipments (NCIMS) with the publication of the Multicriteria-Based Ranking Model for Risk Management of Animal Drug Residues in Milk and Milk Products (the risk-ranking model).

NMPF staff and NMPF cooperative member representatives have supported this work during its multi-year development process and have participated throughout this project in various capacities. The comments offered now represent the culmination of numerous discussions of the published draft report among NMPF staff, our full membership, and dairy industry stakeholders, and in consultation with technical experts on various subjects specific to the risk-ranking model.

NMPF looks forward to continuing our dialogue with FDA on refinement and utilization of the risk-ranking model as one of our tools to develop and implement any potential changes to residue testing requirements in Appendix N of the *Pasteurized Milk Ordinance*.

Respectfully submitted,

A handwritten signature in black ink that reads "Beth Briczinski". The signature is written in a cursive, flowing style.

Beth Briczinski, Ph.D.  
Vice President, Dairy Foods & Nutrition

*The National Milk Producers Federation (NMPF), based in Arlington, VA, develops and carries out policies that advance the well-being of dairy producers and the cooperatives they own. The members of NMPF's cooperatives produce the majority of the U.S. milk supply, making NMPF the voice of more than 32,000 dairy producers on Capitol Hill and with government agencies. Visit [www.nmpf.org](http://www.nmpf.org) for more information.*



---

James Mulhern, *President & Chief Executive Officer* | Randy Mooney, *Chairman*

July 29, 2015

Agri-Mark, Inc.  
Associated Milk Producers Inc.  
Bongards' Creameries  
Cooperative Milk Producers Association  
Cortland Bulk Milk Producers Cooperative  
Dairy Farmers of America, Inc.  
Dairymen's Marketing Cooperative, Inc.  
Ellsworth Cooperative Creamery  
Farmers Cooperative Creamery  
FarmFirst Dairy Cooperative  
First District Association  
Foremost Farms USA  
Land O'Lakes, Inc.  
Lone Star Milk Producers  
Maryland & Virginia Milk Producers Cooperative Association  
Michigan Milk Producers Association  
Mid-West Dairymen's Company  
Mount Joy Farmers Cooperative Association  
Northwest Dairy Association  
Oneida-Madison Milk Producers Cooperative Association  
Prairie Farms Dairy, Inc.  
Premier Milk Inc.  
Scioto County Cooperative Milk Producers' Association  
Select Milk Producers, Inc.  
Southeast Milk, Inc.  
St. Albans Cooperative Creamery, Inc.  
Swiss Valley Farms Company  
Tillamook County Creamery Association  
United Dairymen of Arizona  
Upstate Niagara Cooperative, Inc.  
Zia Milk Producers, Inc.

Division of Dockets Management  
HFA-305  
Food and Drug Administration  
5630 Fishers Lane, Room 1061  
Rockville, MD 20852

**Re: Comments on Multicriteria-Based Ranking Model for Risk Management of Animal Drug Residues in Milk and Milk Products; Docket Number: FDA-2015-N-1305**

Dear Sir or Madam:

The National Milk Producers Federation appreciates the opportunity to comment on the Multicriteria-Based Ranking Model for Risk Management of Animal Drug Residues in Milk and Milk Products (the risk-ranking model). The National Milk Producers Federation (NMPF), based in Arlington, VA, develops and carries out policies that advance the well-being of dairy producers and the cooperatives they own. The members of NMPF's cooperatives produce the majority of the US milk supply, making NMPF the voice of more than 32,000 dairy producers on Capitol Hill and with government agencies. Visit [www.nmpf.org](http://www.nmpf.org) for more information.

NMPF supports the ongoing efforts of the federal Food and Drug Administration (FDA) to improve the safety of the US raw milk supply and processed milk and dairy products, including the development of the risk-ranking model. NMPF is pleased to see the publication of the risk-ranking model, in fulfillment of a 2008 request from the National Conference on Interstate Milk Shipments (NCIMS), recognizing the significant resources that were directed toward this project. The risk-ranking model will be a valuable tool to aid in making science-based evaluations of current and future milk testing programs. NMPF respectfully offers the following comments related to the approach of the current model, the assumptions and scientific data used, and clarity and transparency of the model and report, as well as its potential implications related to drug residue testing programs implemented through the *Pasteurized Milk Ordinance* (PMO) by the NCIMS.

**Criterion A: Likelihood of Drug Administration to Lactating Dairy Cows.**

Criterion A (likelihood of drug administration to lactating dairy cows) consists of four distinct sub-criteria to score the probability a particular drug will be used to treat a lactating dairy cow, a dry cow or a heifer: published surveys and a formal expert elicitation panel, marketing status of the drug, approval status of the drug, and farm inspection data. While there are limitations of the data, the general approach does seem reasonable.

**Sub-criterion A1: Likelihood of drug administration score based on published surveys and formal expert elicitation.**

The report identified the most influential sub-criterion for ranking drugs in Criterion A to be A1. Therefore, NMPF feels compelled to strongly emphasize the need for accurate and representative data in A1's factors (the Sundlof survey, the USDA survey and the expert elicitation panel). NMPF offers the following four recommendations for further consideration:

1. The Sundlof national veterinarian survey should be removed from the model.
2. The assumption that all drugs within the same drug family/group share similar usage patterns, regardless of approval status, should be re-evaluated and appropriately modified.
3. The data from the expert elicitation panel should receive an increased weight.
4. The expert elicitation panel should be expanded to include more participants, as well as to include key subject-matter experts with knowledge of drugs for which information is currently lacking.

The rationale for these four recommendations is described in greater detail below.

The report discusses limitations of the national veterinarian survey (Sundlof) conducted in 1992 – including timeliness of the data. NMPF agrees that data collected twenty-three years ago is highly questionable, as drug approval status and availability, animal health and well-being, and farm management practices have significantly changed during this time period. Additionally, the Sundlof data was limited in that data on all 54 drugs in the model was not available, and assumptions were made within drug families/groups. **This suggests the Sundlof survey is no longer representative of the current dairy industry and is of limited value for the current work and, absent a repeat of this national survey, should be removed from the risk-ranking model.** If this dataset remains in the model, then it should receive significantly less weight than the other factors within Sub-criterion A1.

There are also limitations of the survey data collected by USDA, through the National Animal Health Monitoring Survey (NAHMS), specifically that the survey does not collect data on individual drugs, but across drug classes. With both the NAHMS survey and the Sundlof

survey, the assumption was made that drugs in the same drug class have the same likelihood of being used if they are used to treat the same conditions. While this may be a reasonable assumption to make if all the drugs within a group or class share the same approval status, it is extremely unlikely that two drugs within the same class have similar usage patterns when their cost of use, approval, or marketing status differs – the amphenicols (florfenicol versus chloramphenicol), sulfonamides (sulfabromomethazine and sulfadimethoxine versus sulfaquinoxaline and sulfachloropyridazine), anti-parasitics (eprinomectin and moxidectin versus amprolium), and NSAIDs (flunixin versus phenylbutazone) as prime examples. While the approval status of a drug is taken into account later in the model (Sub-criterion A3), this assumption likely results in a significant over-estimation (increased score for Sub-criterion A1) of use for drugs that are not approved for lactating dairy cows or food-producing animals, yet are within a family that does contain a drug that is either approved for use in lactating dairy cows or may be used under FDA's extra-label drug use policy. **This assumption (for Factors A1.1 and A1.2) should be re-evaluated in families with drugs of differing approval statuses and adjusted accordingly.**

An expert elicitation panel was convened to reduce the bias inherent in Factors A1.1 and A1.2 (the Sundlof veterinarian data and the USDA NAHMS survey) and to provide additional information on individual drugs. NMPF agrees with the need to use an expert panel to provide additional information on likelihood of drug use in lactating dairy cows and offers two additional recommendations to improve the accuracy of this factor.

**First, the expert panel should receive an increased weighting within Sub-criterion A1.** This would acknowledge the limitations, assumptions, and biases in the two surveys (noted above) as well as would recognize that the information from the panel provides drug-specific information that reflects current industry practices – which is otherwise lacking from the Sundlof and USDA datasets in their present form. While our first preference, as recommended above, is to remove the Sundlof data from the model entirely until a similar survey is repeated, if this factor remains in the model, then the expert panel should receive more weight in recognition of its increased accuracy and relevance to Sub-criterion A1.

**Second, NMPF suggests expanding the number of participants in the expert elicitation panel that was used to estimate likelihood of drug use in lactating dairy cows (Factor A1.3), as well as the likelihood of the drug's presence in milk (Sub-criterion B3).** A panel of nine experts was asked to estimate:

- the percentage of dairy herds to which a drug is administered during a year,
- the percentage of dairy cows within herds to which a drug is administered during a year,
- the average number of treatments per year,
- the likelihood of a drug entering a cow's milk after administration, and
- the likelihood of contaminated milk entering the bulk-milk tank.

Expanding the number of panelists who participated in the expert panel would have two key advantages. First, inherent limitations of using an expert panel would be minimized (individual biases, geographical differences, etc.). Second, those who are true subject-matter experts in more specialized fields can be afforded the opportunity to participate and provide much-needed information to the model. For complete coverage of the five questions asked of the expert elicitation panel, both pharmacologists (with knowledge of drug metabolism in dairy cows) and active veterinarian practitioners (with knowledge of drug use on dairy farms) would need to be involved. While each group contributes a unique and valuable perspective, it is unlikely that any single panelist is able to contribute accurate information across all five questions for all 54 drugs.

In fact, the panelists themselves recognized this knowledge gap by the large proportion of “no responses” that were recorded. “No response” was provided for between 15% of the drugs (the first three questions, for Sub-criterion A1) and more than 20% of drugs (the last two questions, for Sub-criterion B3) by at least two-thirds of the panelists. Clearly the expert elicitation panel, as assembled in its current form, was not able to provide the information desired for completeness and accuracy of the model.

Seven of the drugs (~13% of the 54) received a “no response” from at least two-thirds of the panelists for all five of the questions posed (clorsulon, oxfendazole, sulfabromomethazine, sulfaethoxypridazine, sulfaquinoxaline, thiabendazole, tilidipirosin), which suggests the panel’s input could be improved by seeking out those with knowledge of these specific drugs. Two of these drugs appeared among the top 20 drugs in the final overall ranking (#11: oxfendazole and #14: sulfaquinoxaline) and an additional three appeared among the top 25 (#21: thiabendazole; #22: sulfaethoxypridazine; and #24: sulfabromomethazine). Assuring accurate information about the likelihood of use and the likelihood of residues occurring in milk for all drugs examined in the model is critical to the accuracy of the ranking overall. **More reliable data can be obtained by expanding the number of participants in the panel, as well as seeking out subject-matter experts with specific knowledge of drugs for which “no response” was commonly provided.**

**Sub-criterion A4: Likelihood of drug administration score based on evidence of the drug’s presence on dairy farms, based on farm inspection data.**

Sub-criterion A4 determines a score for each of the 54 drugs based on FDA inspection reports, related to the number of times each drug was identified on a dairy farm during an FDA dairy farm inspection between 2008 and 2014. The model assumes the presence of a drug on a farm to be correlated to likelihood of that drug being administered to a lactating dairy cow.

NMPF points out that these farms were specifically targeted for inspection – as a result of a tissue residue violation in an animal presented for slaughter at a processing plant – and should not be considered “typical”. In fact, dairy farms linked to cull cow tissue residue violations are a definite minority. In 2011, there were 453 tissue residue violations in cull dairy cows (per FDA data). USDA-NASS reports there were 51,300 licensed dairy farms in the United States during that time, meaning less than 0.88% of dairy farms were implicated in a tissue residue violation. These violations represent individual drugs identified at levels in excess of US tolerances, not the number of individual animals (i.e., a single animal could have a violative level of penicillin and a violative level of flunixin in tissue, which would count as two separate violations), nor does it indicate if multiple violations were from the same dairy farm, so the overall percentage of dairy farms having a tissue residue violation is likely even less than 0.88%.

In the absence of other survey information about what drugs may be found on a dairy farm (and therefore have an increased likelihood of use), the survey data may be a reasonable proxy for estimating likelihood of use for some drugs, especially for those that are approved for lactating dairy cows. However, as tissue residue violators are an atypical population, this data likely over-estimates the likelihood of use for drugs that are not approved for lactating dairy cows. **NMPF seeks additional clarification in the report that the dairy farms which were inspected are not typical of most dairy farms because of the unique characteristics of the population surveyed, as well as the potential implications of using this data.**

As an alternate source of on-farm survey data, NMPF recalls a dataset based on inspections conducted by Milk Sanitation Rating Officers (this was part of a discussion of the risk-ranking model previewed by the NCIMS Appendix N Modification Committee). The inspection data was collected by FDA during ratings. These data, representing a statistically valid sample, would be more representative of drugs found on-site for the entire population of dairy farms compared to inspection data based on farms of tissue residue violators. NMPF questions if the check rating inspection data were, in fact, replaced in the model with the tissue residue violator inspection data, and asks FDA to re-evaluate the value of the data from the check rating inspections. **NMPF’s recommendation is to use the checking rating inspection data as the basis for Sub-criterion A4, rather than the data representing an atypical sub-population.** Should FDA not agree with removing the tissue residue violator inspection data, then NMPF would suggest the score from Sub-criterion A4 be weighted significantly less than the scores from the NAHMS survey and expert panel (A1) and the drug marketing status (A2).

Additionally, NMPF noted an inconsistency in scoring Sub-criterion A4. The scoring system reported in the text (Table 5.10, p 33) is based on percentages of farms (i.e., <1%, ≤10%, ≤30%, ≤45%, >45%). However, the scoring system reported in Appendix 5.2 for Sub-criterion A4 is based on an absolute number of inspections for which a drug was identified

(i.e., 0-1, >1, >10, >50, >150). For ease of discussion, the two scoring systems have been consolidated into the table below:

<b>Drug Scoring For Sub-criterion A4</b>		
	<b>Text, p 33</b>	<b>Appendix 5.2, p 158</b>
<b>Score</b>	<b>Drug identified in X of farms inspected</b>	<b>Drug identified in X inspections</b>
9	>45%	>150
7	>30% to ≤45%	>50
5	>10% to ≤30%	>10
3	≥1% to ≤10%	>1
1	<1%	0-1

In comparing the two tables, it appears that the percentages and numbers of inspections are not equivalent. For example, if a particular drug was noted on more than 10 farms during inspections, the drug would score a 5 (per the Appendix); the text notes that drugs scoring a 5 represent at least 10% of the total inspections. Together, this would suggest that there were about 100 farms (10% of 100 farms is 10 farms) in the total dataset. However, per the scoring system in the Appendix, drugs scoring a 9 were identified on more than 150 farms (a larger number of farms than the estimated total). In fact, Appendix 5.7 (Table A5.16) indicates the total number of farms in the dataset was 979 farms.

Perhaps an easier way explain the inconsistencies observed between the two scoring systems is to compare the two scoring systems in the text and in the Appendix through their relative differences in score values. For example, drugs scoring a 3 versus a 5 in both the text and the Appendix differ in their prevalence on farms by a factor of 10 (e.g., 1% versus 10%). However, drugs scoring a 5 versus a 7 differ by a factor of 3 in the text (10% versus 30%) and differ by a factor of 5 in the Appendix (10 farms versus 50 farms). **NMPF recommends this inconsistency be resolved, as it appears there are two different scoring systems in the report, and seeks clarification as to which scoring system was actually used to generate the output of this risk-ranking. Further, NMPF suggests reporting scores relative to percentages may be preferable in the future, provided there is a footnote as to the total number of inspections contained in the dataset as well as a reference to the source of the data (i.e., type of inspections, time period).**

**Weighting of Sub-criteria within Criterion A.**

Within Criterion A, an expert elicitation panel suggested increasing the weight of Sub-criteria A1, A2, and A4 (survey data, drug marketing status, and evidence of drug use on farms) relative to A3 (drug approval status). While this approach is reasonable, **NMPF suggests slightly increasing the weight of Sub-criterion A1 (survey data)** because, according to the report, the likelihood of drug administration “*LODA score based on published surveys*

*and formal expert elicitation... is most directly relevant to the question at hand*" (p 25), as well as A1 being *"the most influential sub-criterion for ranking drugs in criterion A"* (p 64). Increasing the weight of A1 relative to the weights of A2 and A4 would appropriately reflect this conclusion about the relevance and usefulness of the data. Also, as suggested above, **the expert panel (Factor A1.3) should receive an increased weighting within Sub-criterion A1**, as this data is the most relevant, drug-specific, and reflective of current industry practices compared to the two other datasets (Sundlof and USDA NAHMS surveys).

**Criterion B: Likelihood of the Drug's Presence in Milk (Bulk-tank or Bulk Milk Pickup Tanker).**

Criterion B (likelihood of drug's presence in milk (bulk-tank or bulk milk pickup tanker), given that the drug is administered to lactating dairy cows) consists of three distinct sub-criteria: sampling data from the National Milk Drug Residue Database and FDA CVM Milk Drug Residue Sampling Survey, potential for misuse of the drug, and a formal expert elicitation panel. As with Criterion A, there are a number of limitations of the data, and data is lacking on multiple drugs among the 54 represented in the model.

In general, the inclusion of Criterion B in the model offers valuable information and the approach is reasonable. Together with the score from Criterion A, the likelihood score from Criterion B contributes to an estimation of the probability of a particular drug residue occurring in the milk supply. NMPF respectfully offers the following comments on the specific sub-criteria, factors and data that were used to inform an overall score for Criterion B.

**Sub-criterion B1: Score for likelihood of drug presence based on evidence that the drug has been identified in milk (bulk-tank milk or bulk milk pickup tanker).**

Sub-criterion B1 is based on evidence – the National Milk Drug Residue Database (NMDRD) and the FDA sampling survey (CVM survey) – that a particular drug has been identified in milk. If a drug was identified in milk through the NMDRD (Factor B1.1), it was assigned a 9; if a drug class was identified in milk, it was assigned a 7; and if a drug was not identified, it was assigned a 3. **NMPF disagrees with the conclusion that a drug that was not actually identified in milk, but is a member of a drug class for which a member had been identified, be assigned a significantly higher score (7 versus 3).** As mentioned earlier, broad generalizations across a drug family or class may be reasonable if all drugs within that family share similar characteristics (approval/marketing status, etc.). While this assumption may be reasonable within some families (e.g., beta-lactams), it is extremely unlikely that this would apply across families when the drugs are known to vary widely in their cost of use, marketing, or approval (and, hence, likelihood of use) status. Examples include the



amphenicols (florfenicol versus chloramphenicol), sulfonamides (sulfabromomethazine and sulfadimethoxine versus sulfaquinoxaline and sulfachloropyridazine), anti-parasitics (eprinomectin and moxidectin versus amprolium), and NSAIDs (flunixin versus phenylbutazone). This assumption is likely to result in an overestimation of the likelihood score for the presence of some drugs in milk – especially for drugs not approved for use in food-producing animals – without supporting data as to true prevalence. Approval and marketing status are considered elsewhere in the model, but would not sufficiently offset the increased score a drug may receive under Sub-criterion B1 based on the current assumption and structure of the model.

Likewise, the assumption also depends upon the specific test kit identified in the NMDRD having similar sensitivities for all drugs within a particular class or family. For some families (e.g., beta-lactams), this assumption may be reasonable. For other drug families, the test kit may not detect all drugs within that family, or may not detect all drugs within that family at sensitivities that are relevant. As an example, the test kit reported in the NMDRD for the aminoglycoside family detects gentamicin, neomycin, streptomycin and dihydrostreptomycin at or below US safe or tolerance levels (detection <100 ppb). The test will also detect amikacin and kanamycin, but at levels that are fairly high (>1,000 ppb). This is an ideal example of when it is not reasonable to conclude an equal likelihood of a particular test kit detecting all drugs within that family. Because these two drugs are the only two aminoglycosides evaluated in the study which do not have formula approved in food-producing animals (the other aminoglycosides do have at least one formula approved for use in cows, per the formula approval status in Appendix 3.2, p 137), along with their high detection level by the test kit reported to the NMDRD, it is not appropriate to apply the same elevated score across all aminoglycosides.

Drugs that have been identified in the milk supply have a greater probability of entering the milk supply in the future than drugs not identified in the milk supply, and the scoring should reflect as such, without extending the assumption to encompass all drugs within all classes or families. It is not reasonable to suggest that a drug receive an elevated score indicating a greater likelihood of prevalence in bulk tank milk in the absence of any data to suggest this is true. **NMPF recommends this broad assumption across all families of drugs relative to their presence in milk be removed from the model – or limited only to those drug families where such an assumption is reasonable based on actual patterns of drug usage and residue prevalence, as well as test kit capabilities.** It is not clear from the report if this incorrect assumption also applies to Factor B1.2 (the FDA milk drug residue survey). NMPF requests this be clarified and, if such an assumption was made with respect to this factor, that it be removed from this portion of the model as well.

NMPF is also concerned that the evidence used to develop a score for Sub-criterion B1 is based on the identification of the drug in milk. The model only considers whether or not a

drug has been identified in milk, without considering prevalence. According to the model, if a single violation in milk was identified, then the drug would score a 9. If the drug has never had a single violation reported, then it would score a 3 (with the assumption described in the paragraphs above removed from the model). By default, the score for Sub-criterion B1 is the maximum score a drug receives from either Factor B1.1 (the NMDRD survey) or Factor B1.2 (the FDA milk residue sampling survey). With this scoring system, a single violation – regardless of data source or total number of tests performed – truly results in a maximum score.

Residue testing is performed for a number of reasons – general or routine testing (such as might be done to meet regulatory requirements, customer specifications, or other business-to-business agreements) as well as risk-based testing (such as might be done when a producer suspects they may have accidentally added residue-containing milk to a bulk tank). The current approach is likely to result in an overestimation of a drug's presence in milk by treating all drug violations to be of equal likelihood of occurrence. By ignoring the prevalence of a drug residue in milk, residues that represent outliers are scored with equal likelihood as those drugs that more commonly occur, yielding an inaccurate snapshot of residues across the entire milk supply. Such an approach does not seem reasonable, nor would it add as much value to the model as would prevalence data.

Granted, there are limitations to scoring the drugs identified in the NMDRD based on prevalence. It has been commonly recognized that, while the third-party database captures residue violations or positive test results, there is no reliable estimation of “a denominator”, or the total number of tests that were performed for a particular drug. Despite this limitation, the NMDRD does seem to be the most reliable source of publicly available data from which prevalence could be estimated. While it may result in an over-estimation of the true prevalence of many drugs reported to the database, as additional testing data is accumulated over time, the risk-ranking model can easily be updated to reflect a more accurate estimation of prevalence. Therefore, **NMPF suggests Sub-criterion B1 be revised to include an estimation of prevalence of a drug being present in milk (bulk-milk tanker).**

With the recommendation to convert Sub-criterion B1 from a score based on absolute identification to a score based on relative prevalence, NMPF also questions why this sub-criterion is based on a scale of 3 to 9, rather than a scale of 1 to 9. **NMPF recommends anchoring the scale at both 1 and 9, or seeks clarification as to why the scores would not be anchored by the lowest point on the scale when there is no data to suggest likelihood of the drug being present.**

It was noted that the data from the National Milk Drug Residue Database was based on reports for fiscal years 2000 to 2013. This fourteen-year span could be modified to represent simply the last ten years for which data is currently available (fiscal years 2005

through 2014), and updated to the most recent ten years each time the model is re-run. According to the NMDRD, the percent of positive tankers in fiscal year 2014 was 0.014%, which is a significant decrease from a positive rate of 0.073% in fiscal year 2000. While there has been a significant reduction in the positive rate for beta-lactams between 2000 and 2013 (~79%), there has also been a similar reduction in the positive rate for non-beta-lactam drugs during this same time period (~91%). These decreases (more than 80% for all drugs reported) suggests the NMDRD data from fourteen years ago may no longer be representative of current on-farm management practices with respect to use of drugs and likelihood of drug residue prevalence in bulk-tank milk. **NMPF suggests the data from the NMDRD that is used in the model be limited to the last ten years (or less) for which data is available (currently 2005-2014), and that the data be updated each time the model is re-run.**

This recommendation to limit the NMDRD data used in the model to the most recent decade will not substantially change the drugs for which data is available. While ceftiofur, chloramphenicol, gentamicin, and sulfachloropyridazine were only specifically tested for between 2000 and 2004 (years which would be omitted by switching to NMPF's recommended dataset), there were no positives reported to the database; therefore, there would be no change to the scoring for those drugs. Sulfathiazole and chlortetracycline were only specifically tested for between 2000 and 2004 as well, but were not included among the 54 drugs in the model (and no positives were reported). **The changes in the dairy production landscape (number of farms, size of farms, location of farms, etc.) and changes in on-farm management practices would suggest it is reasonable to limit the NMDRD data to the most recent ten years (or less), especially if the NMDRD data is to be used to determine prevalence of residues in milk.**

One of the data gaps identified in the report was *"additional milk testing data to more comprehensively and quantitatively estimate the prevalence and level of each of the 54 drugs and related metabolites in bulk tank milk"* (p 72). Additional testing data does exist, but may not be available publicly (e.g., through the third-party database) for various reasons – for example, some state regulatory agencies may not report non-beta-lactam test results to the third-party database, while other state regulatory agencies only report positive results (thereby not giving an accurate "denominator" for total number of tests performed). If a formal request were made (i.e., through a *Federal Register* notice), industry trade associations (e.g., NMPF) could provide a potential means for broadly gathering testing data that is not currently being reported to the third-party database; however, there is little incentive to do so with the model in its current form whereby a single violation results in a maximum score for likelihood of residue prevalence. Working collaboratively with industry to gather additional testing data (of both positives and total numbers of tests performed) would benefit the model in allowing for a more accurate estimation of prevalence.

With respect to Factor B1.2 (the CVM milk drug residue survey), it is important to note the purpose of this study, which was to determine “if dairy farms with a previous tissue residue violation have more drug residues in raw milk than other dairy farms” (FDA CVM report, 2015). The study was conducted by collecting milk samples from dairy farms with a previous tissue residue violation (targeted farms) and from a comparable number of randomly selected dairy farms that were not selected for inclusion in the targeted list (non-targeted farms).

While the statistical analysis ultimately demonstrated there was no significant increased risk of violative residues in dairy farms from the targeted group, FDA noted qualitative differences between the two groups, which were discussed in greater detail in that final report. In summary, florfenicol (a drug included among the 54 drugs in the risk-ranking model) was found in milk samples from both targeted and non-targeted groups, but only samples from the targeted group contained additional residues of four other drugs.

As emphasized throughout the milk drug residue survey report by FDA, only the targeted group contained drug residues other than florfenicol, which suggests the targeted group is unique in some respect to the general population of dairy farms (represented by the non-targeted group in this survey) for the types of drugs for which residues are likely to be identified in milk. **Therefore, while NMPF recommends turning to the NMDRD for prevalence data in the model, the same recommendation would not be made for the FDA CVM drug residue sampling survey.** It would be inaccurate to draw conclusions about prevalence from the FDA CVM drug residue sampling survey when the survey results indicated a clear qualitative difference between the targeted group and the industry as a whole. At the very minimum, any use of the data from the FDA CVM drug residue sampling survey should be limited to that of the non-targeted group.

**Scoring Sub-criterion B1: LODP based on evidence that the drug has been identified in milk (bulk-tank milk or bulk milk pickup tanker).**

Scoring of Sub-criterion B1 is based on examining the scores for a drug for Factor B1.1 (NMDRD) and Factor B1.2 (FDA milk sampling survey), and using whichever score is numerically greater (p 37 of the report). Drugs that were specifically identified through the NMDRD or the FDA CVM milk residue sampling survey scored a 9 for the factor, and those drugs that were not identified scored a 3. Drugs that were among a class identified in the NMDRD scored a 7 and positive non-violative results from the FDA CVM survey received a score of 5 for Factors B1.1 and B1.2, respectively.

The data for the 54 drugs from the NMDRD appears in Table A5.18 (Appendix 5.8, p 180) and the data for the drugs identified by the FDA CVM milk sampling survey appears in Table A5.19 (Appendix 5.9, p 182). NMPF’s understanding of the scoring system is such that the

data from these tables was converted into Factor B1.1 and B1.2 scores for each of the 99 formulations as shown in Appendix 6.2, Figure A6.4 (p 316).

Hypothetically, a drug scoring a 7 and a 3 for Factors B1.1 and B1.2, respectively, would receive a 7 for Sub-criterion B1 (the higher of the two scores). Likewise, a drug scoring a 3 for both Factors B1.1 and B1.2 would receive a 3 for Sub-criterion B1. However, this does not seem to be the scoring system which is presented in the report. For example, Ivermectin-2 and Tripelennamine received a 3 for both Factor B1.1 and Factor B1.2, yet scored a 9 for Sub-criterion B1. Sulfamethazine-3 received a 9 for both Factors B1.1 and B1.2, yet scored a 3 for Sub-criterion B1.

It is also not clear how different formulations of the same drug also received different scores based on the NMDRD and survey data available (which identifies the presence of a drug or drug family, not specific drug formulations). For example, Penicillin G Procaine-3 and Penicillin G Procaine-2 both received a 7 and 3 for Factors B1.1 and B1.2, respectively. NMPF would have expected this drug to receive a score of a 7 for Sub-criterion B1; however, formulation 3 scored a 9 for Sub-criterion B1, while formulation 2 scored a 3. **NMPF requests additional clarification as to how scores were calculated for Factors B1.1 and B1.2, as well as Sub-criterion B1. The scores displayed for the 99 formulations in Figure A6.4 do not correctly match NMPF's expectations for at least 66% of the formulations.**

**Sub-criterion B3: Score for likelihood of drug presence based on expert elicited information.**

Sub-criterion B3 also made use of a panel of experts to provide input on the likelihood of a drug entering a cow's milk after administration and the likelihood of the drug (in the udder milk) entering the bulk-tank or bulk milk pickup tanker. As with Criterion A, use of an expert elicitation panel is supported by NMPF because it will provide much-needed information that is drug-specific (rather than attributed broadly across an entire drug class or family).

The concerns raised in comments above (under Sub-criterion A1) with respect to use of this particular panel of experts also apply to Sub-criterion B3. Specifically, "No response" was provided for 22% of drugs for at least two-thirds of the panelists, indicating a significant lack of information. **The expert elicitation panel should be expanded to include more participants, as well as to include key subject-matter experts with knowledge of drugs for which information is currently lacking.** The required information on these drugs is undoubtedly available, it is just a matter of identifying the appropriate experts and seeking their participation.

**Criterion C: Relative Exposure to Drug Residues in Milk and Milk Products.**

The third criterion in the model (C: relative exposure to drug residues from consumption) appears to be the most scientifically complex in terms of the data which was taken into consideration (product composition, impact of heat, etc.), yet also seems to be the criterion that is most lacking in specific information for each of the drugs included in the risk ranking. While we will offer a more general comment on incorporation of Criterion C into the model later in our comments, we also submit the following on its Sub-criterion (C1 and C2) and factors.

**Factor C1.1: Product-composition value.**

The relative hydrophobicity of each drug is mentioned as an important factor in justifying Factor C1.1 (Product composition value); however, **better and more complete data is warranted for this key factor in the model.** The product composition value (C1.1) is based on the relative fat content of each of the products, referencing a pharmacology textbook (Pandit 2011) and solubility in octanol compared to the solubility in water, which is not necessarily applicable to the complex matrices represented by specific dairy products. While pH effects are noted in the report, binding to proteins and other components is not considered. It is important to understand the binding/distribution of the drugs *in vivo* (through typical product manufacture) as opposed to when simply added to a test tube of milk or to a “model” (non-dairy) system.

Limited information is available for some of the drugs on partitioning in milk and milk products (Table A5.24). Whether there is enough data to be useful is open to question, as only nine drugs were examined in cream/milk, and only eight drugs were examined in cheese/milk. This is a significant data gap and is critical in determining a score for this sub-criterion. **NMPF strongly urges additional data be developed to improve the accuracy of scoring for Factor C1.1.**

It should further be noted that data on partitioning of drugs between milk and milk products is not solely dependent upon fat solubility. When comparing relative concentration of residues in a dairy product versus the concentration in milk, this is also dependent upon binding to other components, ionic environment, partitioning between phases, pH, time, and processing steps involved in product manufacture. Therefore, the data used to calculate grades for Factor C1.1 (partitioning in milk, cream, cheese) are not completely independent of what would be reflected by heat degradation and water removal values (Factors C1.2 and C1.3). **NMPF questions how this overlap could be addressed in refining the model in the future.**

**Factor C1.3: Water removal value.**

The model takes into account the impact of “water removal” (Factor C1.3) via condensing or evaporation, which assumes the concentration of drug residues would increase. However, this factor does not take into account water that may be removed through removal of whey (as through cheese manufacture) or through selective filtration and the fate of residues as a result of those processes. Whey proteins are efficient at binding hydrophobic compounds (especially  $\beta$ -lactoglobulin<sup>1</sup>), and any such residues that may be bound by whey proteins ultimately may not be concentrated in the final dairy product. As mentioned above, this is a limitation of Factor C1.1 (product composition) by only considering partitioning behavior between lipid and water. For both cheese manufacture and selective filtration, **the extent of binding of drugs to casein and whey proteins should be included among the factors for Sub-criterion C1.**

Filtration is currently an important processing procedure that is expected to grow in importance in the future, and there are multiple types of filtration (nano-, ultra-, micro-, and dia-) that are selective in the fractions that they make. However, filtration is not represented among the processing steps considered in Sub-criterion C1. Inclusion of this processing step becomes even more important when one considers recent improvements and new ingredients based on this technology that are becoming more common in the marketplace (e.g., milk protein concentrates, micellar caseins). Product examples that use filtration technology would include various fractionated whey protein products (WPC-30, WPC-80, WPI) and Greek yogurt, as well as high-solids or high-protein fluid milk products, some of which use multiple filtration steps.

**Filtration technology is not directly incorporated into any of the other proposed dairy products or processing steps, and is an additional process that should be considered.**

Condensing or evaporation would concentrate antibiotics in the heated, retained portion and thus would not necessarily reflect the effect of lower temperatures and molecular size filtration and the impact of protein binding during processing that would occur through filtration. Similarly, cheese manufacture would not be a good substitute for filtration in that pH, ionic environment, protein concentration (including which specific proteins) and fat concentration effects on the partitioning of antibiotics would be important variables in cheese manufacture that are not reflective of filtration.

**Sub-criterion C2: Magnitude of consumption of dairy products.**

There are multiple ways to determine the consumption patterns of dairy products. Each of the ways has strengths and weaknesses that need to be considered before use in the

---

<sup>1</sup> Papiz, M, L Sawyer, E Eiopoulos, A North, J Findlay, R Sivaprasadarao, T Jones, M Newcomer, and P Kraulis. 1986. The structure of  $\beta$ -lactoglobulin and its similarity to plasma retinol binding protein. *Nature* 324: 383-385.

current risk-ranking model. The What We Eat in America/National Health and Nutrition Examination Survey (WWEIA/NHANES), which was used in the current model, was originally designed to approximate the usual nutrient intake of Americans age 2 and older. For this approximation, the evaluators use the Automated Multiple-Pass Method to help participants recall the food they consumed within the last 24 hours on two separate occasions. There have been a number of peer-reviewed journal articles mentioning the limitations of this approach<sup>2,3</sup>, while recognizing that a major strength of these data is the large number of participants, which contributes power to the data analysis.

**An alternate means of estimating consumption of dairy products would be USDA’s Dairy Product Per Capita Consumption, United States, Annual Report<sup>4</sup>**, from which the following data was obtained:

Product	Pounds consumed per capita
Milk and Cream	189.0
Yogurt	14.9
Butter	5.5
American Cheese	13.4
Other (mostly Mozzarella)	20.1
Cottage Cheese	2.1
Evap/Condensed whole milk bulk and canned	1.9
Evap/Condensed skim milk bulk	5.3
Regular Ice Cream	12.8
Reduced Fat Ice Cream	5.9
Sherbet	1.5
Other Frozen products	1.8
Dry Whole Milk	0.2
Non-fat dry milk	2.9
Dry Buttermilk	0.4
Dry Whey including modified whey products	1.6

There are notable differences between the data in the table above and the data extracted from NHANES. For example, USDA-ERS’s statistics demonstrate mozzarella cheese

<sup>2</sup> Archer, E, G Pavela, CJ Lavie. 2015. The inadmissibility of What We Eat in America and NHANES dietary data in nutrition and obesity research and the scientific formulation of national dietary guidelines. Mayo Clinic Proceedings. 90(7): 911-926.

<sup>3</sup> Auestad, N, JS Hurley, VL Fulgoni III, CM Schweitzer. 2015. Contribution of food groups to energy and nutrient intakes in five developed countries. Nutrients, 7(6): 4593-4618.

<sup>4</sup> USDA-ERS. <http://www.ers.usda.gov/data-products/dairy-data.aspx>, Dairy products: Per capita consumption, United States (Annual), updated 9/2/2014.



consumption to be higher than American (Cheddar and processed) types, which also reflects current US production patterns. The data used in the risk-ranking model (Table 5.30, based on NHANES) indicate that average mozzarella daily consumption over a lifetime is 0.07 g/kg-bw/day versus 0.29 g/kg-bw/day for Cheddar and processed cheeses. The NHANES survey, suggesting mozzarella consumption is less than one-fourth that of Cheddar and processed cheese, is significantly different from USDA-ERS data with mozzarella consumption at 20 pounds per capita versus 13.4 for American. **As the data presented by USDA-ERS also mirrors USDA-NASS production data<sup>5</sup>, these might be additional datasets worth considering.**

Additionally, the latter dataset from NASS does give more detail on the amount of individual products, including whey protein concentrates and lactose, the incorporation of which would add value to the model. Over a long period of time (in this case, estimating over the course of a lifetime), it is reasonable to assume that dairy product production is equivalent to dairy product consumption; hence, USDA-NASS production data could be a reasonable proxy for estimating dairy product consumption in the model.

While NMPF suggests consideration of an alternate dataset (USDA-ERS per capita consumption data or USDA-NASS annual production data), we do recognize this does not take into consideration age- and body weight-adjusted consumption data. However, although the risk ranking model does include age- and body weight-adjusted consumption data, it averages this over a life-time, which somewhat negates the age-adjusted calculation. USDA per capita consumption data could be included in the model, with some conversion into daily intake per kg body weight. **Given the apparent and significant inconsistencies in consumption data between USDA and NHANES datasets, and given that there are more specific data on product consumption and a wider variety of dairy products in the USDA production databases, serious consideration of USDA-derived data is warranted for increased accuracy.**

Notwithstanding the limitations and inaccuracies of the data represented in the NHANES dataset, NMPF recommends the data be checked for accuracy and completeness. Table A5.27 (Appendix 5.16, p 268-305) reportedly contains the dairy products present in foods consumed by those who participated in the WWEIA/NHANES work, with the dairy product, food code, food description, and percent dairy ingredient for each item. However, only the first 16 foods in the table appear to have a percentage of dairy ingredient listed, while the rest of the entries appear to contain none (i.e., “—”). Likewise, many of the food descriptions do not appear to be listed in the table (see p 269-270). **Should the NHANES data remain in the model, NMPF requests the missing data be added for both “Food Description” and “% Dairy Ingredient” to all of the items included in the report.**

<sup>5</sup> USDA-NASS. [http://www.nass.usda.gov/Statistics\\_by\\_Subject/index.php](http://www.nass.usda.gov/Statistics_by_Subject/index.php). Accessed July 1, 2015.

**Identification of Milk and Milk Products for Inclusion in the Model.**

In identifying the specific dairy products to consider in the model, twelve products were selected to represent a diversity of both composition and processing and are commonly consumed. Specifically, protein-enriched dairy ingredients (e.g., WPC, MPC) were excluded, as noted in the report (Appendix 4.1), due to a lack of information on drug binding to milk proteins. As cited in our comments above, this information is both critical and relevant when considering cheese manufacture and other fractionated dairy protein ingredients. If protein-binding data were available, it would be worth **including dry whey and modified whey products (lactose, WPC) among the products in the model**. While NHANES does not include these products among their survey, USDA production data does include these products and indicates their consumption approaches that of non-fat dry milk.

**Another consideration in identifying the specific dairy products in the model would be more accurate reflection of fat content.** Product composition, specifically level of milkfat, is important to determine drug concentration by product. In the current version of the model, production compositions (Table 4.1) reflect general composition data and whole milk or full-fat versions of products, rather than being weighted to indicate current market conditions. For example, current IRI data indicate that whole milk represents 31.5% of the market, and reduced-fat (2% fat) milk, low-fat (1% fat) milk, and fat-free (<0.2% fat) milk represent 37.8%, 18.2%, and 11.9% of the market, respectively. The weighted-average of the fluid milk sales data is 2.0% fat and the most popular form of fluid milk is reduced-fat milk, yet the current model assumes all fluid milk consumed is full-fat (whole milk) at 3.3% fat.

A more striking example is yogurt, where the IRI data indicate that whole milk yogurt is 8.5% of sales, with reduced-fat yogurt, low fat yogurt, and fat-free yogurt representing 0.9%, 48.3%, and 42.3% of the market, respectively. In this case, the most popular form is low-fat yogurt (the weighted-average of the yogurt sales data is <0.8% fat); while the model (Factor C1.1) is based on whole milk yogurt (3.3% fat). Additionally, current IRI sales data indicate that Greek-style yogurt, which typically has more protein than traditional yogurt, constitutes at least 36% of all yogurt sold – quite a significant portion of the marketplace.

This speaks to **modifying or weighting the product composition in the model to reflect current consumption patterns, rather than assuming a “traditional” formulation (see Table 4.1)**. More representative fat and protein contents should be reflected in Sub-criterion C1 (through differences in product-composition value). Especially given the increase in scores for hydrophobic drugs and dairy products where the fat has been concentrated, it is important to acknowledge that, for some products, the most commonly consumed form has a lower fat content than that of raw milk, which is not currently taken into account in C1.1. **This again also emphasizes the very critical need for more accurate**

**protein-binding data for dairy products where the protein content and casein:whey protein ratio has been changed from that of fluid milk.**

**Criterion D: Potential for Human Health Hazard.**

Criterion D evaluates the potential for human health hazard, given exposure to a drug residue. NMPF would like to point out an inconsistency in the scoring of this criterion. In the text (Model Description, p 56), the only data used to score Criterion D is the hazard value for the drug. However, in reviewing Appendix 6.3 (p 342) “Animal Drug Data Confidence Score for Criterion D”, the report states that the drug-related data used in Criterion D include 1) the hazard value and 2) the carcinogenicity of the drug. **NMPF requests this inconsistency be resolved and clarified in the final report.**

**General Comments – Formulations versus Drugs.**

FDA selected 54 veterinary drugs and their various formulations for evaluation. For Criterion A (likelihood of drug’s administration) and Criterion B (likelihood of drug’s presence in bulk-tank milk), it was necessary to consider specific formulations of each drug separately, as they had unique attributes that would impact their scores for those criteria (e.g., Sundlof survey, FDA inspection data, approval status, marketing status, withdrawal time).

The information for these formulations was used to determine the overall scores for Criteria A and B for each of the 54 drugs; however, it is not clear how differences in scores among formulations of the same drug were resolved. For example, with respect to the various sub-criteria and factors under Criteria A and B, the scoring for the three formulations of penicillin G procaine differed, yet were ultimately resolved into a single score for “penicillin” for each criterion. **NMPF requests clarification in the report as to how a single score for a drug was determined when there were differences in scores among the drug’s formulations.**

**NMPF also seeks further explanation as to the specific formulations that were considered for each of the 54 drugs.** While the text repeatedly refers to 99 formulations that were evaluated through the current work, numerous tables (see Appendices 3.2, 5.5, 5.6 and 5.7) only contain data for 98 formulations. This inconsistency should be addressed and clarified throughout the report.

### **General Comments – Inclusion of Criterion C in the Model.**

The NCIMS Executive Board requested that FDA conduct a risk analysis, including a series of questions from the NCIMS Drug Residue Committee (p 102). One of the questions posed for inclusion in the risk analysis was the influence of dairy product processing on drug residues found in bulk tank milk: *Of the potential drug residues found in bulk tank milk, what is the fate of these residues during processing/manufacturing of various milk products (that is where and at what concentrations would these residues be found in milk products)?* This question was based on the fact that there is limited data on the detection of animal drug residues in processed dairy products or of actual reported cases of illness associated with animal drugs in processed dairy products.

FDA addressed this question through Criterion C, relative exposure to drug residues in milk and milk products, which consisted of two sub-criteria (impact of processing on drug residue concentrations and magnitude of consumption of dairy products). **After extensive discussion and thoughtful consideration, NMPF suggests that Criterion C be removed from the model for the following reasons:**

1. The data that contributes to the scoring of the various factors and sub-criteria of Criterion C are extremely limited. NMPF has provided detailed comments above on the limitations of the existing data, the key data gaps and need for more accurate information, and the additional points that should be considered within this portion of the model. This part of the model, more than any other criterion, would benefit from additional science-based, drug-specific information. However, in its current form, the limited information and assumptions within this criterion minimize the accuracy of the scores obtained for the 54 drugs, without providing reasonable separation (differentiation among the drug scores for Criterion C).
2. The NCIMS Drug Residue Committee requested FDA examine the impact of processing on drug residues with the idea that perhaps processing negated the presence of certain drugs from the food supply. The current data would suggest this is not true.

The values assigned to heat inactivation levels (C1.2) appeared to have little impact on scoring compared to values assigned to partitioning of drugs into the fat (C1.1). In fact, the report acknowledges that the two drug classes ranked highest in terms of relative exposure were due to their hydrophobicity. In general, the heat inactivation factor might be considered insignificant in the model.

Ultimately, Criterion C scoring was based on a relative exposure value (C1 x C2). A score of 5 was assigned for Criterion C if the final relative exposure value was  $\leq 6.0$  and a score of 9 was assigned if the value was  $> 6.0$ . Reducing a complex criterion with multiple sub-criteria and factors to a bimodal distribution, one that is primarily dependent upon

hydrophobicity, reflects the incomplete nature of the available data, which would otherwise have been able to effectively separate the drugs in terms of the relative impact of processing over the entire scale (1 through 9). The fact that the ultimate scoring of Criterion C reflects the most lipophilic drugs only (Figure A6.11, p 325) suggests the data used to develop Criterion C is incomplete. As the impact of processing appears to be minimal (from what information is available), the lack of a significant effect would also support removal of Criterion C from the model at this time.

While the impact of processing and magnitude of consumption are important in assessing potential risks to consumers from animal drug residues in the food supply through a risk assessment or risk analysis, the current model only provides a risk-ranking of the 54 drugs in relation to each other, without providing a true indication, or quantitative assessment, of the magnitude of human health concerns. The NCIMS Appendix N Program is a compliance program, to test for drug residues in Grade “A” raw milk supplies. Based on this premise, the most critical criteria in the model are related to the likelihood of animal drugs being used in lactating dairy cows and being present in the bulk tank milk (i.e., Criteria A and B).

Based on the significant data gaps, as well as the apparent lack of significant impact presented by Criterion C, NMPF questions whether it is appropriate that Criterion C remain in the model. While the full model (including Criterion C) addresses the presence of drug residues in the food supply, the NCIMS residue testing program is a compliance program for the raw milk supply. **NMPF suggests serious consideration be given to removing this criterion from the model entirely at this time.** A ranking of drugs without Criterion C would also allow for education and outreach programs targeting those drugs that are of greatest likelihood for actually occurring in raw milk.

#### **General Comments – Avermectins.**

The high scores and rank for many of the antiparasitic drugs were derived from a combination of high and higher than average scores for all four criteria. NMPF would like to respectfully call attention to a few points which may have incorrectly, and repeatedly, elevated the scores throughout the model for this class of drugs – ultimately affecting the final ranking of this drug class.

- Avermectins can be applied to dairy cows via topical or injectable routes and those that are approved in lactating dairy cows (eprinomectin, moxidectin) have a “zero” milk withholding time, which is a *de facto* recognition that these drugs will not be present in milk at levels of human health concern. This has been supported by the FDA milk sampling survey, where only a single sample of milk contained a violative concentration of an anthelmintic (doramectin). Despite their typical mode of application and zero milk

withholding time, the avermectins typically scored a 5 overall for Criterion B (likelihood of presence of the drug in the bulk-tank milk), which seems to be over-estimated for the avermectins approved in lactating dairy cows.

- As was discussed previously with Criterion C, the bimodal scoring of the criterion (based on a cut-off score of 6), the over-simplification of drug partitioning in dairy products (in non-dairy systems without protein binding), and the weighting of dairy products of higher fat content than is actually consumed (not incorporating lower fat products into Factor C1.1) would all contribute to a higher score for the hydrophobic anti-parasitic drugs than would be reasonably expected.
- The antiparasitics class is another example of where the assumption that all drugs within a class share similar likelihoods of use, misuse, and residue prevalence is unrealistic. Within the antiparasitics, there are avermectins, oral dewormers, an anti-coccidia drug (amprolium), and an anti-fluke drug (clorsulon). These drugs differ in their application, their approval status, their marketing status, their cost, etc. Throughout the model, specifically with Criteria A and B, drugs within a class are scored similarly. This is a prime example of a drug class where such an assumption likely results in the over-estimation of scores for many of the drugs within the class, especially those that are not often used in dairy cattle (e.g., amprolium, clorsulon, and thiabendazole).

Taken together, these assumptions and limitations of the data likely over-estimate the final score and ranking of this class of drugs. **NMPF suggests that assumptions related to this class of drugs be re-evaluated, additional information be sought to more clearly differentiate drugs within the family, and alternative ways of scoring the drugs be considered.**

#### **General Comments on Implications and Use of the Model.**

In the Executive Summary, Results, Answers to Charge Questions, and Conclusion sections, the report repeatedly calls attention to the “top 20” highest-scoring drugs. In fact, FDA proposed testing for the drugs in eight different drug classes which ranked among the “top 20” highest-scoring drugs in a proposal at a recent meeting of the National Conference on Interstate Milk Shipments. NMPF has supported a risk assessment process as a tool to make science-based decisions about how drugs should be re-evaluated for inclusion in milk testing programs. However, **NMPF does not understand the logic or rationale behind focusing on a “top 20”** (or “top 15”, “top 10”, or “top 3”, etc.) among a list of 54 drugs, which seems arbitrary.

Additionally, FDA's recommendation is not limited to the "top 20" drugs which are the highest-scoring, rather the eight different drug classes ranked among the top 20 highest-scoring drugs. This recommendation is overly broad. In fact, those eight drug classes (beta-lactams, antiparasitics, macrolides, aminoglycosides, NSAIDs, sulfonamides, tetracyclines, amphenicols) represent 45 of the 54 drugs, or 83% of those on the list. NMPF suggests that conducting a risk ranking, only to ultimately end with a recommendation that encompasses the vast majority of the drugs considered, runs counter to the purpose of developing a risk ranking which is to identify those specific drugs of particular public health concern which should be considered for inclusion in milk testing programs. By extending the recommendation from individual drugs to entire drug families, FDA has essentially given all drugs in the same class the highest "priority score" of the entire class. **NMPF recommends that FDA revise the recommendation to focus on individual drugs rather than drug classes.** This will allow for more effective focusing of resources within milk testing programs and within targeted residue prevention materials (i.e., education, outreach programs).

FDA has made public statements about how the model will be used to inform decisions related to milk testing programs (i.e., Appendix N of the *Pasteurized Milk Ordinance*). Essentially, FDA has suggested that as new data is gathered, the model would be re-run to generate a new ranking of drugs, and that testing programs would continue to focus on the classes represented by classes of the new "top 20" highest-scoring drugs.

NMPF raises two questions with this approach:

- 1) How will FDA determine when to re-run the model? Will it be at a pre-determined interval (e.g., every two years before each biennial NCIMS meeting), will it be based on when new data is available (e.g., an update to farm inspection data, a new NMDRD report, changes in drug market or approval status), or will it be as data gaps are able to be filled (e.g., availability of experimental data on the fate of drug residues as a result of processing)? Will the list of 54 drugs be re-evaluated each time the model is re-run? How often will the expert elicitation panels be re-convened, if at all, to assure the data reflects the most current patterns of drug use on dairy farms? **NMPF recommends the model be periodically re-evaluated and its input data updated such that the risk-ranking model is representative of current industry practices and, especially, most accurately estimates the likelihood of drug administration and the likelihood of prevalence of drug residues in bulk-tank milk (Criteria A and B).**
- 2) In addition to NMPF's comments above related to a broad recommendation that extends from specific drugs to general drug classes and the "top 20" being an arbitrary cut-off, NMPF also is concerned with this approach that basically amounts to a periodic "re-shuffling of the deck".

Always requiring the drug classes represented by the “top 20” to be tested does not take into consideration any risk management options that may be implemented as a result of the risk-ranking. A basic principle of food safety is that you can’t test your way to safety, but targeted strategies to mitigate or reduce the risk of a particular hazard can be adopted and can be effective at controlling risk. If the dairy industry and/or stakeholders were to implement mitigation efforts to pro-actively manage or reduce the risk of a particular drug residue (evidenced through changes in likelihood of drug use or prevalence of residues in bulk-tank milk), presumably this would be reflected in a reduction in its overall criterion score in the model. **It may be reasonable to consider providing an incentive (for example, through establishment of a threshold overall score) such that if the overall risk represented by a drug or the overall score of a drug is reduced, there would be a similar reduction in the recommendation to test for that drug.**

By always having to test for the drugs in the “top 20” – or the classes represented by the “top 20” – it is unlikely that the drugs of concern will ever change (currently 83% of the drugs on the list), and the costs of testing will continually increase, imposing a burden on the industry without justification of an actual public health or farm compliance issue. It is better to promote implementation of strategies to control for the presence of drug residues in the milk supply than to try to test to assure the absence of residues. The recommendation from FDA for the drugs (or drug classes) of concern should be revised to reflect this.

### **Conclusion.**

The model was designed as a “risk ranking” for the 54 drugs selected for evaluation, rather than a quantitative public health “risk assessment” of drug residues in milk that might inform or warrant changes to mandatory milk testing requirements of Appendix N. As a ranking tool, the current model provides useful information on relative risk among the drugs examined, but lacks utility in evaluating the level of public health risk for even the highest ranked drugs identified by the multicriteria-based decision analysis. Accordingly, NMPF urges FDA to interpret the output of the model appropriately, and to use the information it generates as one, but not the only, tool for guiding the development and implementation of any potential changes to Appendix N requirements.

NMPF recognizes the considerable work by FDA in undertaking development of the risk-ranking model and commends those involved for their thoughtful and methodical approach. We appreciate FDA’s consideration of the comments offered above (as well as additional items for consideration, please see Appendix 1 attached) and hopes this feedback will contribute toward greater accuracy, clarity, and utility of the risk-ranking model.



NMPF looks forward to continuing to work together with FDA on this project in a spirit of cooperation and collaboration. Please contact us if you have any questions.

Respectfully submitted,

A handwritten signature in black ink that reads "Beth Panko Briczinski". The signature is written in a cursive style with a large initial 'B'.

Beth Briczinski, Ph.D.  
Vice President, Dairy Foods & Nutrition

**Enc: Appendix 1, Additional Items for Consideration**

**Appendix 1**  
**Additional Items for Consideration**

NMPF truly appreciates the significant resources that went into developing the risk-ranking model. Our main comments primarily focused on potential suggestions to improve the overall model design and approach, as well as the various sub-criteria and factors, data, and assumptions. During the course of our review of the risk-ranking report, NMPF noted a number of technical edits, which are offered in this Appendix for purposes of accuracy, clarity, and transparency. NMPF requests that FDA consider each suggested technical edit so that the final product from the risk-ranking model reflects the high quality of work which went into developing it.

<b>Location in Risk-Ranking Report (Section, page number)</b>	<b>Suggestion</b>
Abbreviations and Acronyms, v	NAHMS: National (Insert “Animal”) Health Monitoring System
5. Model Description, 19	National (Insert “Animal”) Health Monitoring System (NAHMS)
5. Model Description, 28	...following categories: antibiotics, sulfonamides, anthelminthics... Replace “anthelminthics” with “anthelmintics”
5. Model Description, 33	As noted in comments above, there are two different scoring systems in Table 5.10 (p 33) and Appendix 5.2 (p 158) for Sub-criterion A4. These two scoring systems should be consistent. Viewing scores relative to percentages may be preferable, with a footnote as to the total number of inspections performed.
5. Model Description, 55	Cream (Sour Close parentheses, insert “)”
6. Results, 61	...(especially avermnectins)... Replace “avermnectins” with “avermectins”
6. Results, 62	...the lincosamides: pirlimycin and incomycin... Replace “incomycin” with “lincomycin”

<p>6. Results, 65</p>	<p>...an aminocyclitol (spectinmycin)...</p> <p>Replace “spectinmycin” with “spectinomycin”</p>
<p>6. Results, 70</p>	<p>...(flunixin), and a tetracycline (oxetetracycline)...</p> <p>Replace “oxetetracycline” with “oxytetracycline”</p>
<p>8. References, 74-96</p>	<p>There are a number of inconsistencies with formatting, punctuation, and style throughout the list of references. While some detailed editing to address these is necessary, there are also a number of typos that should be corrected for purposes of clarity. While a more thorough review is suggested, a few examples related to content (not formatting, etc.) are highlighted here:</p> <p>p 74 ...different ripening times following pout-on administration...</p> <p>Replace “pout-on” with “pour-on”</p> <p>p 75 ...food powders – proceses and properties...</p> <p>Replace “proceses” with “processes”</p> <p>p 76 ...FARAD Digest: Extralabel us of...</p> <p>Replace “us” with “use”</p> <p>p 79 ...Committee for Veterinary Medicinal Products for Veterinary Use. Tulathromycin...</p> <p>Delete “for Veterinary Use”</p> <p>p 86 ...FARAD Digest: Extralable use of...</p> <p>Replace “Extralable” with “Extralabel”</p> <p>p 87 ...and Washington Statue University College...</p> <p>Replace “Statue” with “State”</p>
<p>Appendix 3.1, 111</p>	<p>Drugs #13 and #14, laidlomycin and lasalocid, were not included among the final 54 drugs in the study; however, there is no note here indicating why these drugs were removed from consideration.</p>
<p>Appendix 3.1, 132</p>	<p>Title: Table A3.11 Listing of hormones/repro</p> <p>Replace “repro” with “reproductive drug”</p>
<p>Appendix 3.1, 132</p>	<p>Drug #252, drug type: Reporductive/Hormone</p> <p>Replace “Reporductive” with “Reproductive”</p>

Appendix 3.1, 132	Drug #259.1, drug type: hormone Replace “hormone” with “hormone”
Appendix 3.1, 136	Drug #306, no drug type listed.
Appendix 3.1, 136	Drug #312, Prochlorprazine Replace “Prochlorprazine” with “Prochlorperazine”
Appendix 3.2, 137	We note that there are only 98 drug formulas listed, while the text repeatedly mentions 99. This occurs in numerous places throughout the report. The inconsistency should be resolved.
Appendix 3.2, 137	Drug #4.1, Amoxicillin trihydratetrihydrate-1 Replace “trihidratetrihydrate” with “trihydrate”
Appendix 3.2, 142	Drug #52, Tripelemamine Replace “Tripelemamine” with “Tripelennamine”
Appendix 3.2, 142	Footnote [1]... CFR 500-599 (check) Delete “(check)”, or clarify its meaning here.
Appendix 3.2, 142	Footnote [2]... NADA). Is additional information missing? There was a closing parenthesis “)”, but no opening parenthesis “(“.
Appendix 3.2, 142	Footnote [4]... Footnote #4 does not appear in the table anywhere. If it does not need to be cited in the table, then it should be removed here.
Appendix 5.1, 146-155	Tables A5.2, A5.3, A5.4, A5.5, A5.6 Replace “Amprolium” with “Amprolium” Replace “Eprinocectin” with “Eprinomectin” Replace “Sulfachlorpyridazine” with “Sulfachlorpyridazine” Replace “Tripelemamine” with “Tripelennamine”
Appendix 5.2, 158	As noted in comments above, there are two different scoring systems in Table 5.10 (p 33) and Appendix 5.2 for Sub-criterion A4. These two scoring systems should be consistent. Viewing scores relative to percentages may be preferable, with a footnote as to the total number of inspections performed.

Appendix 5.5, p 168	Footnote 29 “Beta-lactam antibiotics are the most widely used group...” The text in the footnote is redundant. It appears both in the footnote and in the second paragraph on page 168. Only one is necessary.
Appendix 5.5, p 170, Appendix 5.6, p 173, and Appendix 5.7, p 176	Amoxicillin tryhydrate-1, (and -2, -3) Replace “tryhydrate” with “trihydrate”
Appendix 5.5, p 170, Appendix 5.6, p 173, and Appendix 5.7, p 176	Ampicillin tryhydrate-1, (and -2, -3) Replace “tryhydrate” with “trihydrate”
Appendix 5.7, p 176	On-farm inspection data for 54 drugs (99 formualtions) Replace “formualtions” with “formulations”
Appendix 5.8, p 179	Grade A bulk-milk pick-p tanker testing Insert quotes around “A” Replace “pick-p” with “pick-up”
Appendix 5.8, p 179	There are two rows of “TETRACYCLINES” in the table. The data in these rows should be summed and should appear in a single row. For clarity, NMPF also suggests consistent capitalization of the drugs within the table.
Appendix 5.9, p 182	The Drug Classes listed in the table do not correspond to the drug classes for which each drug is identified in the report (e.g., Table 3.1, p 11) and should be revised for consistency and clarity.
Appendix 5.9, p 182	Data check for accuracy: Sulfaquinoxaline was assayed for in 191 samples; the other sulfa drugs were assayed for in 1912 samples.
Appendix 5.10, p 183	Ampicillin tryhyrdate-2 is indicated for oral... Replace “tryhyrdate” with “trihydrate”
Appendix 5.10, p 186	Drug #20: ... dosed orally with a capsul... Replace “capsul” with “capsule”

Appendix 5.10, p 186	Drug #22: ... dosed orally with a capsul... Replace “capsul” with “capsule” However, NMPF questions why a study involving detection of nitrofurans is cited as data for an aminoglycoside (gentamicin).
Appendix 5.10, p 192	Drug #52, Formulation Replace “Tripelemamine” with “Tripeleennamine”
Appendix 5.11, p 195	Replace “Pirilomycine” with “Pirlimycin”
Appendix 5.11, p 195	Replace “Tirpeleennamine” with “Tripeleennamine”
Appendix 5.13, p 216	Table A5.24 Replace “Choramphenicol” with “Chloramphenicol”
Appendix 5.13, p 216	Table A5.24. Footnotes All three footnotes have superscript of “a”, instead of “a”, “b”, and “c”.
Appendix 5.14, p 218-263	Table A5.25. Several rows and individual cells in the table appear to be shaded, but not so consistently that the shading appears to have a meaning. Is there a key for shading, or is it random?
Appendix 5.14, p 263	Line numbers may be removed from the margin in the footnote area.
Appendix 5.16, p 268-305	Table A5.27. As noted in comments above, data appears to be incomplete in the table for both “Food Description” and “% Dairy Ingredient”.
Appendix 6.1, p 308	Table A6.1, Criterion A LODA Replace “Acetylsalicyclic” with “Acetylsalicylic”
Appendix 6.1, p 309	Table A6.1, Sub-criterion B2 LODP - Drug misuse Replace “Cehpapirin” with “Cephapirin”
Appendix 6.1, p 309	Table A6.1, Sub-criterion B2 LODP - Drug misuse Replace “Napoxen” with “Naproxen”

<p>Appendix 6.2, p 311</p>	<p>...scores for A1.1, A1.2, and A1.3 (derived from the... Delete open parenthesis, or insert closed parenthesis as appropriate</p>
<p>Appendix 6.2, p 311</p>	<p>... LODA scores in Factors A1.1,-A1.2, and A1.3. Delete negative sign (-) before A1.2.</p>
<p>Appendix 6.2, p 311</p>	<p>A3.... such as phenylbutazone, nitroflurazone, ... Replace “nitroflurazone” with “nitroflurazone”</p>
<p>Appendix 6.2, p 312</p>	<p>Figure A6.1, Figure A6.2, Figure A6.3, Figure A6.4, Figure A6.5 ,Figure A6.6 Replace “tryhydrate” with “trihydrate” for Amoxicillin tryhydrate (1, 2, 3) and for Ampicillin tryhydrate (1, 2, 3)</p>
<p>Appendix 6.2, p 326</p>	<p>Figure A6.12 Replace “Albendazol” with “Albendazole”</p>
<p>Appendix 6.3, p 342</p>	<p>D. Animal Drug Data Confidence Score Drug-related data that are used in Criterion D include (1) hazard value and (2) carcinogenicity. However, the model reports (p 56) only hazard value was used to determine a score for Criterion D.</p>