

Relatório Técnico

Análise de dados do projeto “Ivomec Gold Comprova”

Relatório Técnico- Análise de dados do projeto “Ivomec Gold Comprova”

Este relatório apresenta os resultados das análises dos dados provenientes de experimentos realizados no programa Ivomec Gold Comprova (IGC). As análises, interpretações e conclusões apresentadas estão fundamentadas em conceitos científicos sendo que qualquer esclarecimento de dúvida, críticas ou sugestões devem ser encaminhadas ao autor deste relatório.

Protocolo de Experimento

Os experimentos seguiram protocolo definido pela equipe técnica da Merial, responsável pela condução dos experimentos. O protocolo é adequado, contemplando os principais aspectos de experimentação, tais como aleatorização na definição dos grupos de tratamento, presença de grupo controle, homogeneidade das unidades experimentais no início do experimentos e padronização no sistema de medição e pesagem dos animais.

As análises basearam-se neste protocolo para a modelagem e estimação dos efeitos dos tratamentos e de covariáveis disponíveis no banco de dados.

Consistência dos Dados

O banco de dados original foi fornecido pela empresa Ampla – Consultoria em Informação, responsável pelo desenvolvimento e suporte do software utilizado pela equipe do programa IGC, no formato Microsoft Access.

O banco de dados em seu nível mais baixo apresentou **84749** observações, ou seja, animais e 1190 teste. Este banco foi checado nos seguintes itens:

- Valores nulos nos campos Peso (inicial e/ou final): **39** observações
- Valores “0” nos campos Peso (inicial e/ou final): **4558** observações
- Valores nulos nos campos Data (inicial e/ou final): **0** observações

- Valores “0” nos campos Data (inicial e/ou final): **0** observações
- Tempo entre as pesagens igual a 1 dia: **3170** observações

Estas observações foram excluídas para fins de análise por serem claramente inconsistências na entrada dos dados, totalizando **7767** observações descartadas.

Em posse destas observações partiu-se para a análise da distribuição da variável principal nesta análise, a diferença entre as duas medidas. Alguns testes possuem mais de 2 medidas, e optou-se nesta fase tomar as diferenças entre a primeira medida (menor data de pesagem em um mesmo animal) e a última medida (maior data de pesagem em um mesmo animal). A figura 1 apresenta a distribuição desta variável, chamada de **Dp** por simplicidade.

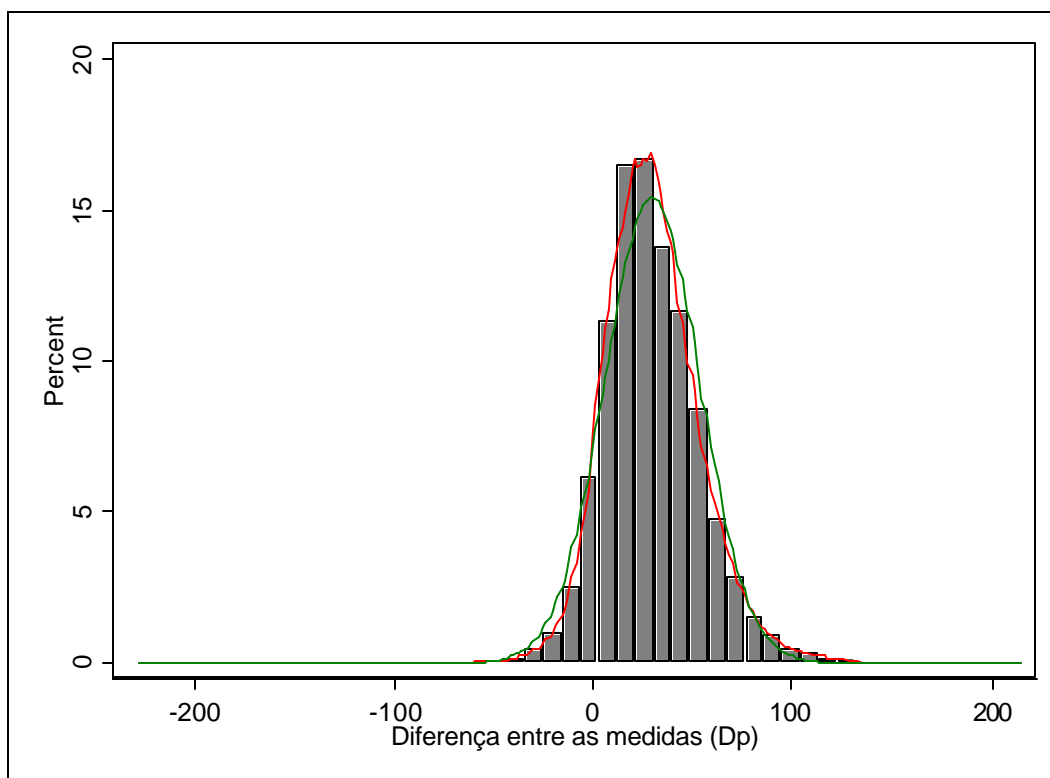


Fig. 1- Distribuição das diferenças de peso (bruta)

Nota-se que a variabilidade é bastante acentuada, com perdas e ganhos de peso chegando a 200 Kg. Tanto do ponto de vista da fidedignidade quanto do

ponto de vista das suposições utilizadas nas análises posteriores, utilizou-se um critério estatístico para se regularizar a distribuição, utilizando-se “trimmed distribution”. A partir de critérios baseados na suposição de Normalidade dos dados (Gaussiana), encontram-se limites mínimos e máximos teóricamente aceitáveis. Os pontos calculados foram

Mínimo: -48 Kg (perda de 48Kg) – 97 observações além do limite

Máximo: 87.5 Kg (ganho de 87.5Kg) - 1327 observações além do limite

As observações que excederam os limites foram descartadas e a nova distribuição apresentada na figura 2.

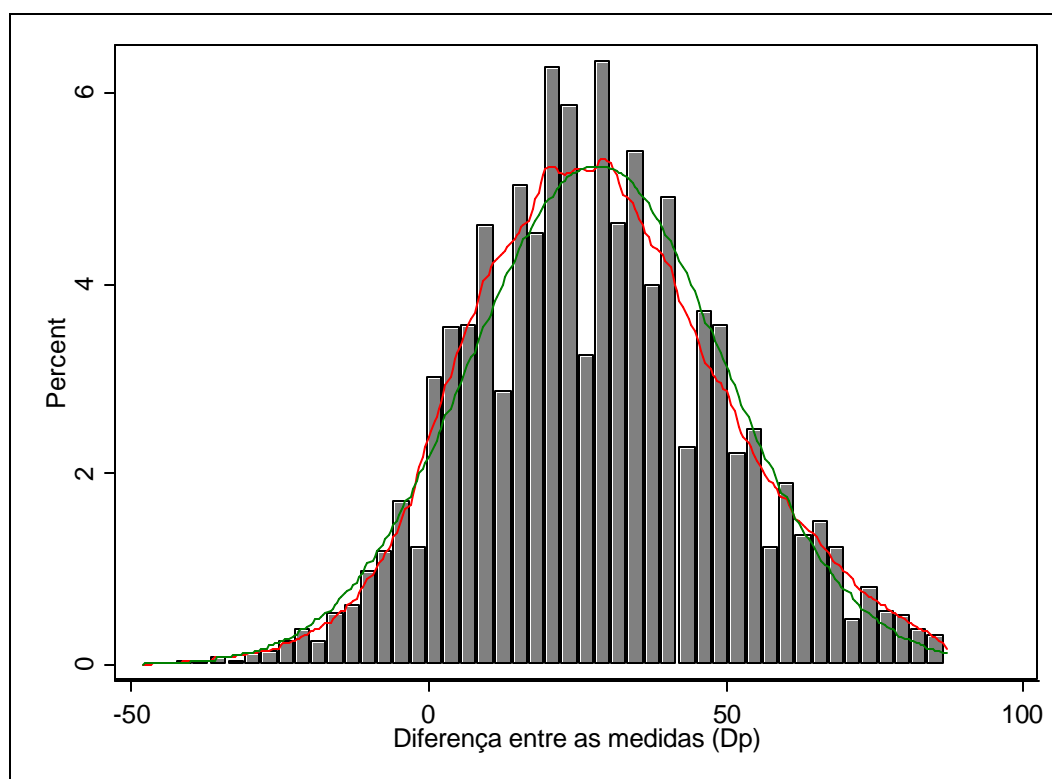


Fig. 2- Distribuição das diferenças de peso (trimmed)

Desta forma a análise será realizada sobre 75558 observações em 1107 testes. O tamanho médio de cada teste foi de 70 animais, sendo um máximo de 120.

Devido às inconsistências e cortes teóricos alguns experimentos tiveram desbalanceamento. Na análise geral destes dados este fator não é agravante, porém pensando-se em análises individuais, técnicas adequadas de análise devem ser empregadas.

Análise dos Dados

A análise de dados seguirá duas abordagens. A primeira, que será chamada de Modelo de Efeitos Aleatórios, levará em conta todos os experimentos em conjunto, como se fosse um “grande experimento” levando em consideração as fontes de variabilidade inerentes desta suposição (propriedade, tempo entre medidas, sinal da diferença, etc...). Uma outra abordagem, de Meta-análise, também será utilizada. A meta-análise leva em consideração os resultados de cada experimento, de forma independente, calculando uma medida de evidência que é uma média ponderada pelas precisões (erros padrão) de cada estudo.

Os detalhes técnicos e referências bibliográficas estão descritos no Anexo Metodológico.

Modelos de Efeitos Aleatórios

A tabela 1 apresenta a distribuição das observações entre os grupos, bem como as médias e desvios padrão da Dp. Observa-se que o balanceamento entre as observações entre os grupos não foi prejudicado. As distribuições de Dp dentro de cada grupo estão apresentadas na Fig. 3, nos Box-plots, onde a linha branca representa a mediana de Dp e os vértices das caixas os percentis 25 e 75.

As análises neste item são realizadas utilizando-se a metodologia de Modelos Lineares de Efeitos Aleatórios, onde o efeito aleatório introduzido refere-se à variabilidade inter-testes, ou seja, foi incorporado ao modelo uma fonte de variabilidade que busca ajustar os efeitos estimados bem como verificar a participação desta na variabilidade total, indicada pelo coeficiente Rho.

Tab. 1- Caracterização dos grupos de estudo para Dp

Grupo	Obs	Freq.	Média Dp	Desv. Pad.
<i>Ivomec</i>	24,678	33.22	32.34	21.04
<i>Controle</i>	24,820	33.41	24.03	21.2
<i>Outros</i>	24,784	33.36	28.4	21.16

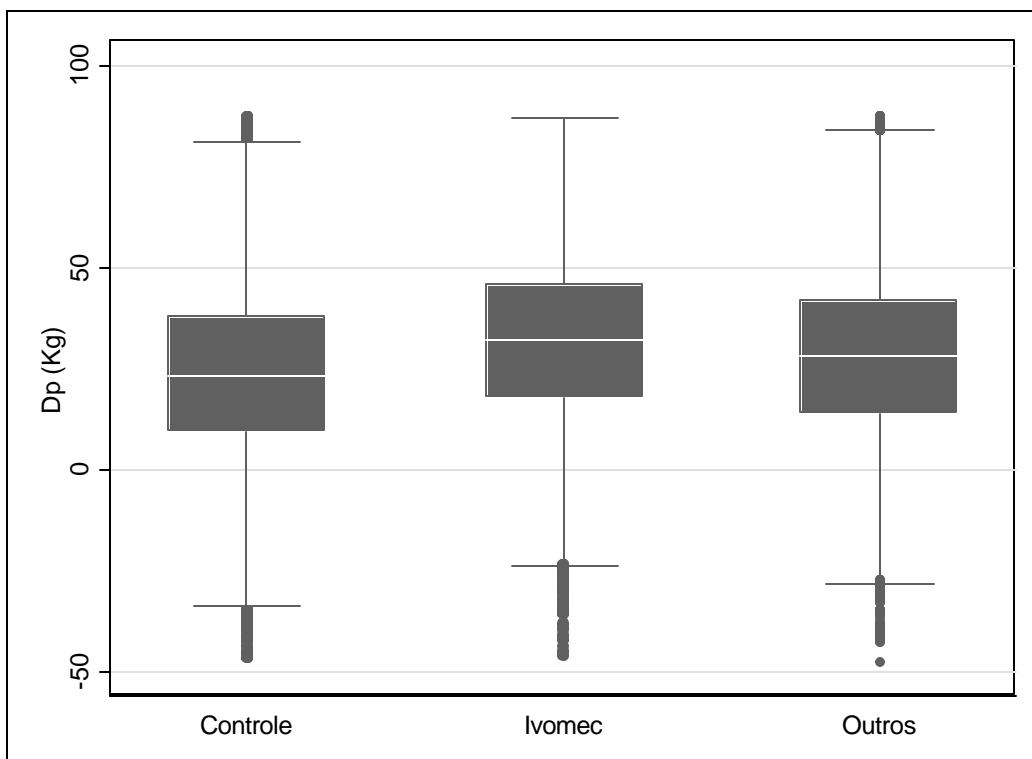


Fig. 3- Distribuição das Dp por grupo estudo

A primeira etapa é verificar se a homogeneidade entre os pesos, ditada pelo protocolo de estudo, foi mantida após os cortes de observações. Nota-se na tabela 3 que estes pesos estão bem homogêneos, com média em torno de 221 Kg.

A tabela 4 apresenta um modelo linear ajustado apenas para a diferença entre grupos. Observa-se que o ganho de peso médio com o Ivomec é de 33.41 Kg, enquanto que observam-se diferenças significativas neste ganho nos outros dois grupos. Comparando-se com o Ivomec o grupo controle apresentou um ganho 8 Kg menor, ou seja, 24.7 Kg, enquanto o grupo tratado com Outros medicamentos apresentou ganho em média 4.1 Kg inferior. As estimativas de cada ganho e seus devidos intervalos de confiança de 95% (IC 95%) estão apresentados na figura 4.

Tab. 3- Pesos iniciais nos 3 grupos

Grupo	Média	ep	IC 95%
Ivomec	221.72	1.72	(218.34, 225.10)
Controle	221.50	1.72	(218.12, 224.88)
Outros	221.27	1.72	(217.89, 224.65)

No entanto observa-se que grande parte da variabilidade se deve às diferenças entre os experimentos, 73%. Buscou-se então caracterizar estas diferenças e verificar sua influência nos efeitos estimados.

O primeiro fator a ser observado é a questão da “direção” das Dp. Alguns experimentos apresentaram perda de peso, enquanto os demais apresentaram ganho. A variável chamada PERDA indica os experimentos onde ocorreu a perda de peso. Incluída no modelo apresentou forte significância e diminuiu o Rho em aproximadamente 10 pontos percentuais. Para simplificar a análise ajustou-se os modelos apenas para experimentos com perda e apenas com ganho de peso, sem “mistura-los”. Observa-se realmente que as magnitudes dos efeitos são diferentes em cada situação. Os valores corrigidos estão apresentados nas Fig. 5a e 5b. Observe que o Rho nos experimentos com perda é bem menor (0.34) em comparação com os de ganho (0.67).

Outro fator que poderia influenciar é o tempo entre as medidas, uma vez que os tempos variam de cerca de 40 dias até 114. Incluído no modelo, o tempo em dias nos experimentos de perda não rejeitou a hipótese de efeito nulo. No entanto, nos experimentos de ganho, o efeito do tempo é de um ganho de *0.27 Kg* por dia (não apresentado). Este efeito é o mesmo no 3 grupos (não há evidência de interação Grupo x Tempo entre Medidas) e não mudou os efeitos de grupo estimados.

Tab. 4- Efeitos estimados pelo modelo

Grupo	Efeito	IC 95%	
Ivomec	33.41	32.32	34.50
Controle	-8.65	-8.84	-8.45
Outros	-4.13	-4.33	-3.94

rho **0.73**

rho: fração da variância devido à variabilidade entre os experimentos

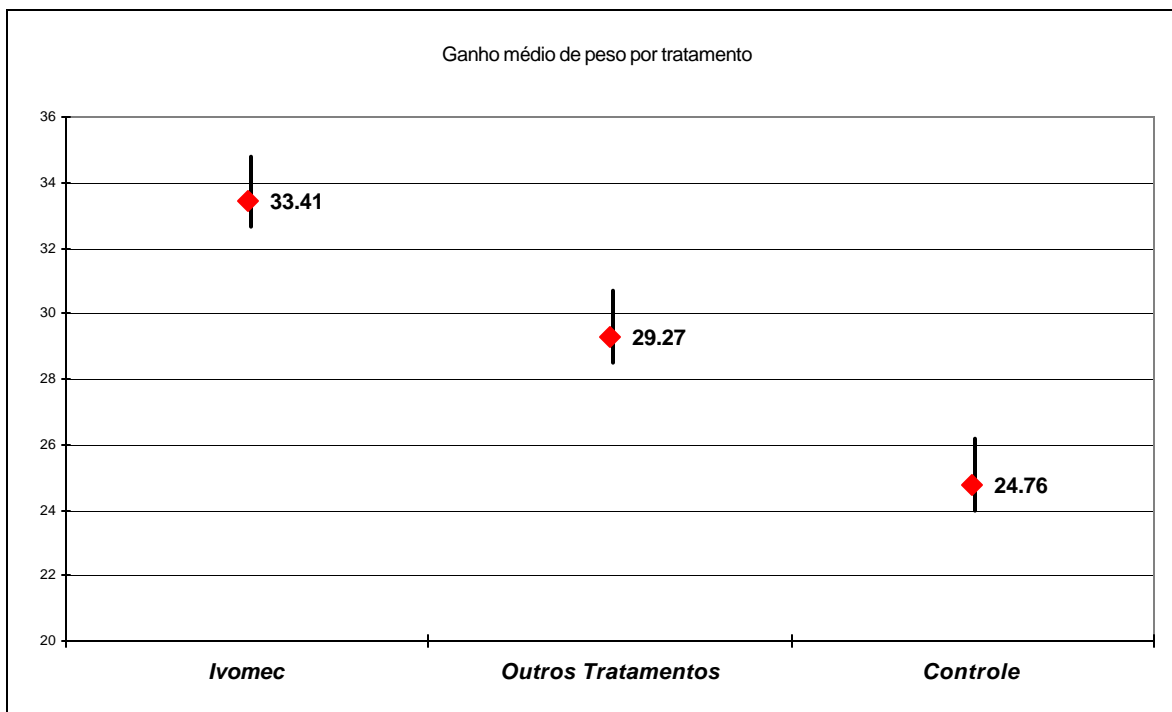


Fig. 4- Médias e IC 95% das Dp nos três grupos de estudo

Tab. 5- Modelos diferenciando estudos com Perda e Ganho de peso

Grupo	Efeito	ep	IC 95%		Fontes de variabilidade	
Perda	-20.29	0.19	-20.66	-19.92	sigma_u	13.84
Controle	-7.48	0.09	-7.66	-7.29	sigma_e	10.42
Outros	-3.64	0.09	-3.83	-3.46	rho	0.64
Ivomec	34.36	0.42	33.53	35.19		
<hr/>						
Experimentos com perda de peso						
Controle	-2.82	0.25	-3.31	-2.33	sigma_u	5.38
Outros	-1.73	0.26	-2.25	-1.22	sigma_e	6.89
Ivomec	-6.76	0.34	-7.43	-6.09	rho	0.38
<hr/>						
Experimentos com ganho de peso						
Controle	-7.80	0.10	-7.99	-7.61	sigma_u	14.86
Outros	-3.71	0.10	-3.90	-3.52	sigma_e	10.42
Ivomec	34.55	0.45	33.66	35.44	rho	0.67

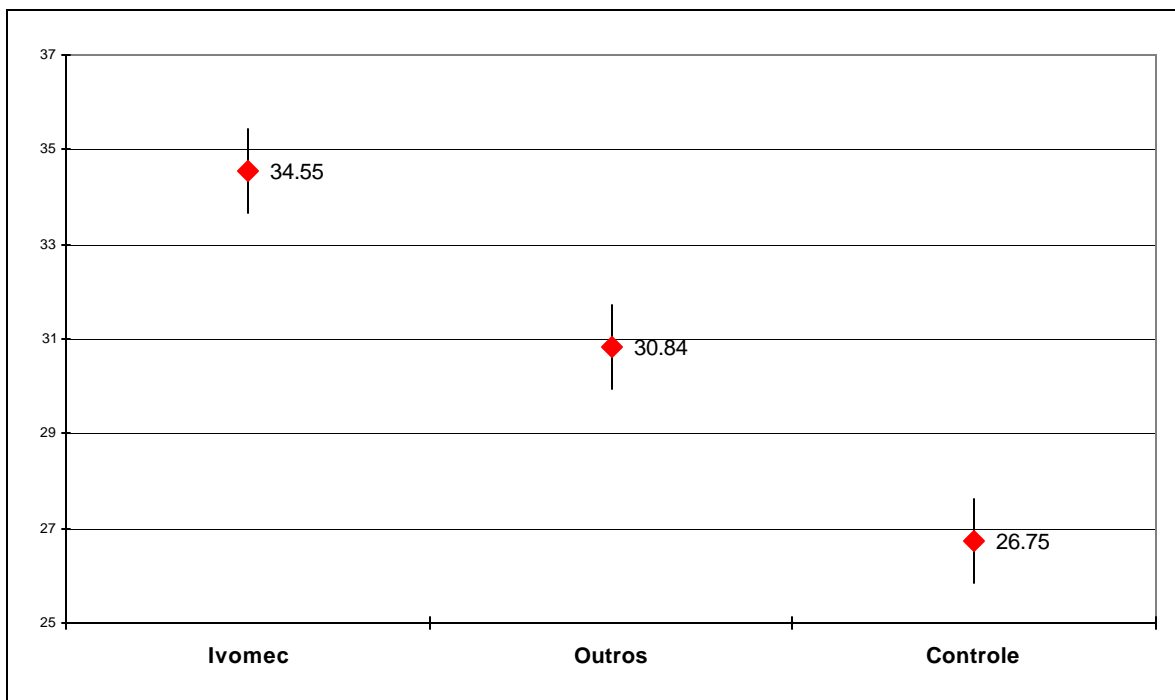


Fig. 5a - Médias e IC 95% nos experimentos com ganho de peso

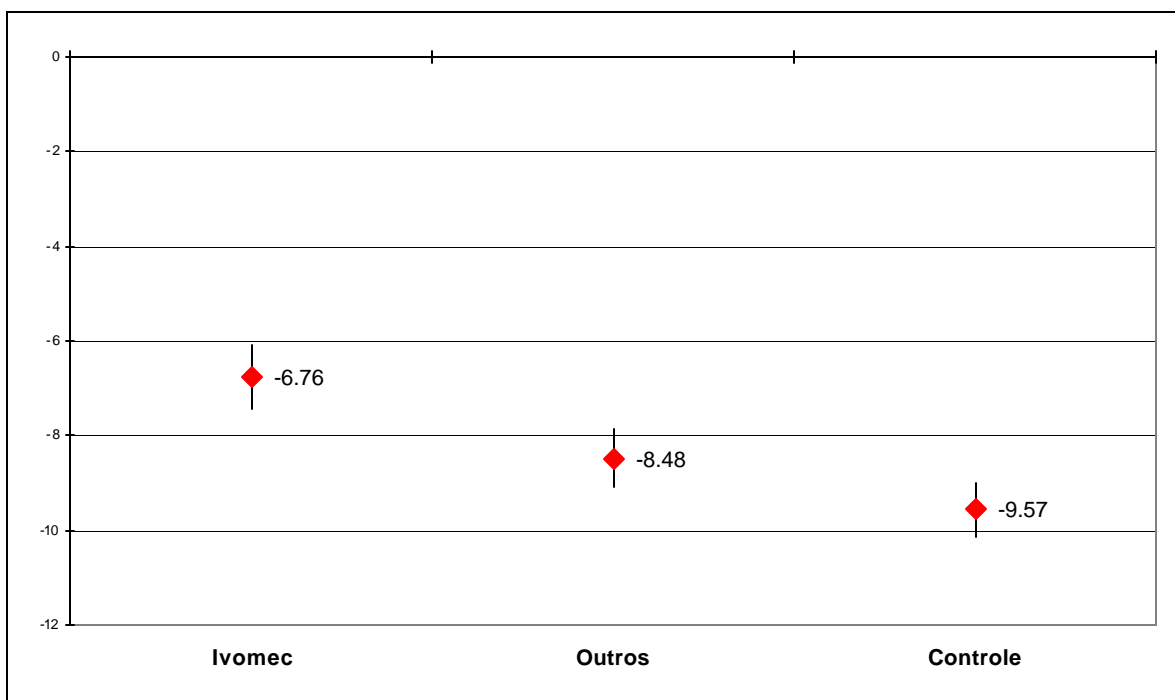


Fig. 5b - Médias e IC 95% nos experimentos com perda de peso

Meta-análise

Os resultados de cada experimento foram obtidos via ANOVA para dados não-balanceados. De cada uma destas análises (1107, uma por experimento) obteve-se o Efeito (diferença em relação à Dp do Ivomec) e os respectivos Erros Padrão, para os dois grupos de comparação, Controle e Outros. A combinação dos resultados também é obtida pelos Modelos de Efeitos Aleatório, no entanto a resposta neste caso é diferente como se pode notar. Seguem os resultados.

Tab. 7- Valores obtidos na Meta-análise

	Efeito	IC 95%	z_value	p_value
	Comb.			
Outros	-4.03	(-4.28, -3.79)	-32.32	0.000
Controle	-8.57	(-8.92, -8.23)	-48.64	0.000

Os resultados apontam que existe forte evidência de que os resultados obtidos com o Ivomec são superiores aos demais grupos. O Efeito Combinado do Ivomec em relação ao grupo Controle é de 8.57 Kg nas diferenças entre os pesos ($p < 0.001$). Em relação ao grupo Outros, este efeito é de 4.03 Kg a favor do Ivomec.

Considerações Finais

As duas abordagens levam à mesma evidência de um ganho de peso maior ou uma perda de peso menor nos animais tratados com o Ivomec Gold. Apesar de parecer redundante, a primeira abordagem é importante para se conhecer as possíveis fontes de variabilidade e a viabilidade da aplicação das técnicas de meta-análise.

A sugestão para implementação computacional é de se desenvolver um aplicativo baseado nos métodos de Meta-análise, conforme descrito no Anexo Metodológico. A implementação da abordagem de Modelos Lineares dentro de cada experimento é válida e é mais uma ferramenta que pode estar disponível aos

técnicos do programa IGC, além das atuais análises descritivas (que são realmente imprescindíveis).

Fernando A. B. Colugnati

Estatístico, MSc

UNIFESP/ EPM

Disciplina de Nutrologia – Dep. De Pediatria

Trabalho desenvolvido em parceria com a Ampla Consultoria em Informação.

Anexo Metodológico

Modelo de Efeitos Aleatórios

O modelo linear utilizado nas análises estima os efeitos principais entre os tratamentos e sua estrutura de variância-covariância incorporando a variabilidade entre os K experimentos como sendo um componente da variância total. Como referência a este tipo de modelo contendo detalhes teóricos sugere-se os livros clássicos de modelos lineares como Graybill (1976). Segue a notação e formulação do modelo.

Seja

y_{it} : Diferença observada entre os pesos final e inicial do i -ésimo animal no t -ésimo experimento;

d_{it}^1 : Variável indicadora do grupo CONTROLE, assumindo valor 1

se o i -ésimo animal no t -ésimo experimento pertence a este grupo, e 0 caso contrário;

d_{it}^2 : Variável indicadora do grupo OUTROS, assumindo valor 1

se o i -ésimo animal no t -ésimo experimento pertence a este grupo, e 0 caso contrário;

o modelo de efeitos aleatórios é definido por

$$y_{it} = \mathbf{m}\mathbf{1} + \mathbf{b}^1 d_{it}^1 + \mathbf{b}^2 d_{it}^2 + \mathbf{X}_{it}\mathbf{d} + u_t + \mathbf{e}_{it}$$

onde \mathbf{m} é a diferença de peso média do grupo tratado com IVOMECA; $\mathbf{1}$ é um vetor unitário de dimensão $N \times K$ (N é número total de unidades experimentais); $\mathbf{b}^1, \mathbf{b}^2$ são os efeitos estimados (diferenças em relação a \mathbf{m}) para os grupos Controle e Outros respectivamente; \mathbf{X}_{it} vetor de covariáveis para a i -ésima unidade no t -ésimo experimento; \mathbf{d} vetor de efeitos das covariáveis; u_t componente da variância entre tratamentos, assumindo distribuição Normal com média 0 e variância σ_u^2 e

e_{it} componente da variância entre unidades experimentais, com distribuição Normal com média 0 e variância s_e^2 .

Os parâmetros de interesse neste trabalho são m e os efeitos de tratamento b^1, b^2 , sendo que seus erros padrão e respectivos testes hipótese são ajustados para os efeitos dos componentes da variância.

A estimação destes parâmetros é realizada pelo método de Quadrados Mínimos Generalizados (Generalized Least Squares). Detalhes sobre este método e procedimentos podem ser encontrados em Wooldridge (2002) e Amemyia (1985).

Nas tabelas apresentadas, os parâmetros estimado σ_u e σ_e referem-se à raiz das variâncias dos componentes apresentados.

Meta-Análise

Esta metodologia encontra larga aplicação nos dias atuais principalmente na área médica, dentro do que se convencionou como Medicina Baseada em Evidências. A evidência seria o efeito combinado de varios experimentos independentes realizados por diferentes grupos ao longo dos anos, testando um mesmo medicamento, sendo que a constatação da eficiência do mesmo é dada por esta medida de evidência. Para maiores detalhes sugere-se a leitura de Hunter & Schmidt (1990).

O efeito combinado pode ser obtido por meio de uma média ponderada dos efeitos de cada estudo, onde os pesos são o inverso da variância destes efeitos. Esta abordagem é conhecida como Modelo de Efeitos Fixos. No caso deste trabalho utilizou-se outra abordagem, chamada de Modelo de Efeitos Aleatórios, ou seja, a mesma apresentada no tópico anterior, sendo que a diferença está na variável de resposta que está sendo modelada. Neste caso o y_{it} não é a diferença entre os pesos final e inicial, mas os β , os efeitos em relação ao grupo tratado com Ivomec. O objetivo final é o cálculo

de um β_p , o efeito combinado, e testa-lo contra a hipótese de que este efeito é igual a zero.

Procedimento Computacional

Neste tópico é apresentado o procedimento computacional para implementação deste tipo de análise.

1. Ajusta-se um modelo de linear para cada experimento utilizando-se o modelo simples

$$y_{it} = \mathbf{m} + \mathbf{b}^1 d_{it}^1 + \mathbf{b}^2 d_{it}^2$$

utilizando-se os estimadores conhecidos de Mínimos Quadrados Ordinários (Graybill, 1976);

2. Os valores de \mathbf{b}^1 , \mathbf{b}^2 e respectivos erros, SE_1 e SE_2 . Estes valores serão utilizados nos cálculos do efeito combinado;
3. Calcula-se o valor da variância entre os K experimentos, para um dado tratamento (e.g. 1= Controle), \mathbf{t}^2 :

$$\text{a) } \mathbf{b}_p^1 = \frac{\sum_{k=1}^K w_k \mathbf{b}_w^1}{\sum_{k=1}^K w_k} \text{ onde } w_k = \frac{1}{SE_1^2};$$

$$\text{b) } Q = \sum w_k (\mathbf{b}_k^1 - \mathbf{b}_p^1)^2;$$

$$\text{c) } S^2 = \frac{\sum_{k=1}^K w_k^2 - K w_p^2}{K - 1} \text{ onde } w_p = \sum_{k=1}^K \frac{w_k}{K};$$

$$\text{d) } U = (K - 1) \left(w_p - \frac{S^2}{K w_p} \right) \text{ e finalmente;}$$

$$\text{e) } \mathbf{t}^2 = \frac{[Q - (K - 1)]}{U}, \text{ se } Q > (K - 1), \text{ ou } 0 \text{ caso contrário.}$$

4. O Efeito Combinado dos K experimentos para um dado tratamento, e.g. 1 = Controle, é dado por

$$\mathbf{b}_p^{1*} = \frac{\sum_{k=1}^K w_k^* \mathbf{b}_w^1}{\sum_{k=1}^K w_k^*}, \text{ onde}$$

$$w_k^* = \frac{1}{SE_1^2 + t^2};$$

5. Obtem-se o IC 95% para \mathbf{b}_p^{1*} com

$$\mathbf{b}_p^{1*} \pm 1.96 SE_{\mathbf{b}_p^{1*}}, \text{ onde } SE_{\mathbf{b}_p^{1*}} = \frac{1}{\sqrt{\sum_{k=1}^K w_k^*}}$$

6. Testa-se $H_0 : \mathbf{b}_p^1 = 0$ com a estatística

$$Z = \frac{\mathbf{b}_p^{1*}}{SE_{\mathbf{b}_p^{1*}}} \sim N(0, 1)$$

Repete-se este procedimento para cada grupo de tratamento, neste caso para os grupos Controle e Outros.

Referências bibliográficas

Graybill, F (1976). *Theory and application of the linear model*. Wadsworth: California

Wooldridge, JM (2002). *Econometric Analysis of Cross-Section and Panel Data*. Cambridge, Massachusetts: MIT Press

Amemyia, T (1985). *Advanced Econometrics*. Cambridge, Massachusetts: Harvard University Press

Hunter, JE & Schmidt, FL (1990). *Methods of meta-analysis: Correcting error and bias in research findings*. Newbury Park, CA: Sage Publications

Sutton, Abrams, Jones, Sheldon S (2000). *Methods for meta-analysis in medical research*. Wiley, Chichester.

Song, Sheldon, Sutton, Abrames, Jones (2001). *Methods for exploring heterogeneity in meta-analysis*. *Evaluation & the Health Professionals*; 24(2):126-51