

MICHERLANIA DA SILVA NASCIMENTO

**ESTIMAÇÃO DA SENSIBILIDADE E ESPECIFICIDADE DE TESTES
DIAGNÓSTICOS PARA A BRUCELOSE BOVINA NA AUSÊNCIA DE PADRÃO
OURO CONSIDERANDO DEPENDÊNCIA CONDICIONAL VIA INFERÊNCIA
BAYESIANA**

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Estatística Aplicada e Biometria, para obtenção do título de *Doctor Scientiae*.

VIÇOSA
MINAS GERAIS - BRASIL
2018

**Ficha catalográfica preparada pela Biblioteca Central da Universidade
Federal de Viçosa - Câmpus Viçosa**

T

N244e
2018

Nascimento, Micherlania da Silva, 1983-
Estimação da sensibilidade e especificidade de testes
diagnósticos para a brucelose bovina na ausência de padrão ouro
considerando dependência condicional via inferência bayesiana /
Micherlania da Silva Nascimento. – Viçosa, MG, 2018.
x, 65 f. : il. ; 29 cm.

Inclui apêndices.

Orientador: Carlos Henrique Osório Silva.

Tese (doutorado) - Universidade Federal de Viçosa.

Referências bibliográficas: f. 51-59.

1. Variáveis latentes. 2. *Brucella abortus*. 3. Testes
imunológicos. I. Universidade Federal de Viçosa. Departamento
de Estatística. Programa de Pós-Graduação em Estatística
Aplicada e Biometria. II. Título.

CDD 22. ed. 519.5

MICHERLANIA DA SILVA NASCIMENTO

**ESTIMAÇÃO DA SENSIBILIDADE E ESPECIFICIDADE DE TESTES
DIAGNÓSTICOS PARA BRUCELOSE BOVINA NA AUSÊNCIA DE
PADRÃO OURO CONSIDERANDO DEPENDÊNCIA CONDICIONAL VIA
INFERÊNCIA BAYESIANA**

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Estatística Aplicada e Biometria, para obtenção do título de *Doctor Scientiae*.

APROVADA: 22 de março de 2018.



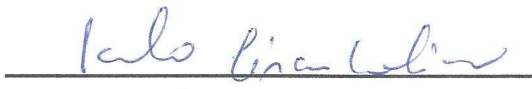
Leandro Roberto de Macêdo



Lidiane Maria Ferraz Rosa



Camila Ferreira Azevedo
(Coorientadora)



Paulo César Emiliano
(Coorientador)



Carlos Henrique Osorio Silva
(Orientador)

*A Deus e a
minha família.
DEDICO*

AGRADECIMENTOS

A Deus, razão da minha vida.

A minha mãe, Cosma Pedro da Silva, pelo amor, amizade, dedicação e incentivo em todos os momentos.

Aos meus irmãos, Rivanio da Silva e Regivaldo Batista da Costa.

À Universidade Federal de Viçosa (UFV) e ao Departamento de Estatística (DET).

Ao meu orientador, Carlos Henrique Osório Silva, por todos os ensinamentos, incentivos e apoio ao longo destes quatro anos.

Ao pesquisador do LANAGRO - MG, Paulo Martins Soares Filho, pelo fornecimento dos dados e disponibilidade, fundamentais para as análises deste trabalho.

Aos meus coorientadores, Camila Ferreira Azevedo e Paulo César Emiliano, pelas importantes contribuições para o melhoramento deste trabalho.

Aos professores, Leandro Roberto de Macedo e Lidiane Maria Ferraz Rosa, membros da banca de defesa, pelas valiosas contribuições.

A todos os professores do Programa de Pós-Graduação em Estatística Aplicada e Biometria.

Aos colegas do curso de doutorado em Estatística Aplicada e Biometria, em especial, a Flávia Sílvia Corrêa Tomaz, pela amizade e palavras de incentivo.

Aos funcionários do Programa de Pós-Graduação em Estatística Aplicada e Biometria, em particular, ao Junior José Pires.

Ao professor Joel Augusto Muniz, da Universidade Federal de Lavras (UFLA), meu orientador de mestrado, reconheço sua importante contribuição durante esta jornada.

Ao professor João Gil de Luna, da Universidade Estadual da Paraíba (UEPB), meu orientador de graduação, por quem tenho um sincero respeito e admiração.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) pelo financiamento da bolsa de estudos.

SUMÁRIO

	Página
LISTA DE FIGURAS	vi
LISTA DE TABELAS	vii
RESUMO	ix
ABSTRACT	x
1 INTRODUÇÃO	1
2 REFERENCIAL TEÓRICO	4
2.1 Brucelose bovina	4
2.1.1 Diagnóstico da brucelose bovina	5
2.2 Medidas de desempenho de testes diagnósticos	6
2.2.1 Estimação com o teste Padrão Ouro	6
2.2.2 Estimação na ausência do teste Padrão Ouro	9
2.3 Modelo de classe latente	11
2.3.1 Independência condicional	13
2.3.2 Identificabilidade	15
2.4 Enfoque Bayesiano	16
2.4.1 Função de verossimilhança	17
2.4.2 Distribuição a priori	18
2.4.3 Distribuição a posteriori	20
2.4.4 Monte Carlo via Cadeias de Markov (MCMC)	21
2.4.4.1 Amostrador de Gibbs	21
2.4.5 Critérios de convergência	21
2.4.5.1 Critério de Gelman e Rubin	22
2.4.5.2 Critério de Raftery e Lewis	23

3	MATERIAL E MÉTODOS	24
3.1	Modelos para testes diagnósticos considerando a suposição de independência condicional e ausência do teste Padrão Ouro	25
3.1.1	Um teste diagnóstico	25
3.1.2	Quatro testes diagnósticos	28
3.2	Modelos para testes diagnósticos considerando dependência condicional e ausência do teste Padrão Ouro	32
3.2.1	Quatro testes diagnósticos	32
3.3	Informação a priori	36
3.4	Crítério de informação da deviance	37
3.5	Procedimentos computacionais	38
4	RESULTADOS E DISCUSSÃO	39
5	CONCLUSÕES	50
6	REFERÊNCIAS BIBLIOGRÁFICAS	51
	APÊNDICES	59

LISTA DE FIGURAS

Figura 1	Gráficos dos traços, correlações e densidades dos parâmetros δ , S e E para o teste diagnóstico AAT.	39
Figura 2	Gráficos dos traços, correlações e densidades dos parâmetros δ , S e E para o teste diagnóstico 2-ME.	39
Figura 3	Gráficos dos traços, correlações e densidades dos parâmetros δ , S e E para o teste diagnóstico FC.	40
Figura 4	Gráficos dos traços, correlações e densidades dos parâmetros δ , S e E para o teste diagnóstico DBac.	40
Figura 5	Gráficos das densidades a posteriori e priori da prevalência da doença (δ) para os testes AAT, 2-ME, FC e DBac.	43
Figura 6	Gráficos das densidades a posteriori e priori das sensibilidades (S) para os testes AAT, 2-ME, FC e DBac.	44
Figura 7	Gráficos das densidades a posteriori e priori das especificidades (E) para os testes AAT, 2-ME, FC e DBac.	44
Figura 8	Gráficos dos traços, correlações e densidades dos parâmetros: prevalência (δ), sensibilidades (S) e especificidades (E) para o Modelo 1.	60
Figura 9	Continuação dos gráficos dos traços, correlações e densidades dos parâmetros: sensibilidades (S) e especificidades (E) para o Modelo 1.	61
Figura 10	Gráficos dos traços, correlações e densidades dos parâmetros: prevalência (δ), sensibilidades (S) e especificidades (E) para o Modelo 2.	61
Figura 11	Continuação dos gráficos dos traços, correlações e densidades dos parâmetros: sensibilidades (S) e especificidades (E) para o Modelo 2.	62

LISTA DE TABELAS

Tabela 1	Resultados de um teste diagnóstico considerando-se a existência do teste Padrão Ouro.	7
Tabela 2	Resultados dos testes diagnósticos AAT, 2-ME, FC e DBac e suas respectivas frequências observadas.	25
Tabela 3	Possíveis resultados observados e latentes para um teste diagnóstico, considerando a ausência do teste Padrão Ouro.	26
Tabela 4	Possíveis resultados observados e latentes para quatro testes diagnósticos, considerando a ausência do teste Padrão Ouro e supondo independência condicional entre os testes.	28
Tabela 5	Possíveis combinações com as respectivas probabilidades para quatro testes diagnósticos, considerando-se ausência do teste Padrão Ouro e supondo-se independência condicional entre os testes.	29
Tabela 6	Estimativas dos hiperparâmetros das distribuições beta para os parâmetros de sensibilidade (S), especificidade (E) para os testes AAT, 2-ME, FC e DBac obtidas conforme os intervalos $[l_1, l_2]\%$	36
Tabela 7	Critérios de convergência de Raftery e Lewis (F_D) e Gelman e Rubin (R).	41
Tabela 8	Médias das distribuições a posteriori dos parâmetros de desempenho dos testes AAT, 2-ME, FC e DBac e da prevalência, com respectivos desvios padrão (DP) e intervalos com 95% de credibilidade (ICr).	42
Tabela 9	Critérios de convergência de Raftery e Lewis (F_D) e Gelman e Rubin (\hat{R}).	45
Tabela 10	Médias das distribuições a posteriori da prevalência, sensibilidades, especificidades e correlações condicionais dos modelos 1, 2, 3, 4 e 5, com respectivos intervalos com 95% de credibilidade (ICr), critério de informação da deviance (DIC) e o número efetivo de parâmetros (p_D).	47
Tabela 11	Critérios de convergência de Raftery e Lewis (F_D) e Gelman e Rubin (\hat{R}).	60

Tabela 12 Desvios padrão das médias das distribuições a posteriori da prevalência, sensibilidades, especificidades, covariâncias e correlações condicionais dos modelos 1, 2, 3, 4 e 5.	63
-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----

RESUMO

NASCIMENTO, Micherlania da Silva, D.Sc., Universidade Federal de Viçosa, março de 2018. **Estimação da sensibilidade e especificidade de testes diagnósticos para a brucelose bovina na ausência de Padrão Ouro considerando dependência condicional via inferência Bayesiana.** Orientador: Carlos Henrique Osório Silva. Coorientadores: Camila Ferreira Azevedo e Paulo César Emiliano.

A brucelose bovina, causada pela bactéria *Brucella Abortus*, é uma doença presente em todas as regiões do Brasil e provoca elevados prejuízos econômicos. O Programa Nacional de Controle e Erradicação de Brucelose e Tuberculose Animal (PNCEBT) estabeleceu os testes AAT, 2-ME, FC e DBac para realizar o diagnóstico da brucelose bovina. Na ausência de um teste Padrão Ouro, é necessário que o desempenho desses testes diagnósticos seja validado. O presente estudo, teve como objetivo empregar o modelo de classe latente Bayesiano para estimar as sensibilidades e as especificidades dos testes diagnósticos AAT, 2-ME, FC e DBac, aplicados em amostras de sangue e carcaças de animais suspeitos de brucelose bovina, bem como a prevalência da doença. O conjunto de dados utilizado foi obtido junto ao Laboratório Nacional Agropecuário de Minas Gerais (LANAGRO-MG). Os testes foram avaliados em dois cenários: individualmente e combinados. Os modelos para a avaliação dos testes combinados foram ajustados considerando-se a independência condicional entre os quatro testes e também incorporando-se ao modelo a dependência condicional entre os testes AAT, 2-ME e FC. As análises foram realizadas em R 3.2.5 usando o pacote *R2OpenBUGS*. Quanto à avaliação dos testes combinados, os resultados mostraram que os testes AAT, 2-ME e FC são condicionalmente independentes. O teste FC foi o mais sensível, o DBac o menos sensível e os testes AAT, FC e DBac foram os mais específicos. Concluiu-se que nenhum dos quatro testes pode ser utilizado sozinho para o diagnóstico da brucelose bovina. Uma baixa sensibilidade foi encontrada para o teste AAT, resultado que diverge dos relatos geralmente encontrados na literatura. Portanto, recomenda-se que contínuos estudos sejam realizados para que a tomada de decisão dos pesquisadores não seja comprometida. Adicionalmente, concluiu-se que o modelo de classe latente bayesiano permitiu estimar os parâmetros de interesse satisfatoriamente.

ABSTRACT

NASCIMENTO, Micherlania da Silva, D.Sc., Universidade Federal de Viçosa, March, 2018. **Estimation of the sensitivity and specificity of diagnostic tests for bovine brucellosis in the absence of the Gold Standard considering conditional dependence via Bayesian Approach.** Adviser: Carlos Henrique Osório Silva. Co-advisers: Camila Ferreira Azevedo and Paulo César Emiliano.

Bovine brucellosis, caused by the bacterium *Brucella Abortus*, is a disease present in all regions of Brazil and causes high economic losses. The National Program for the Control and Eradication of Brucellosis and Animal Tuberculosis (PNCEBT) established the AAT, 2-ME, FC and DBac tests for the diagnosis of bovine brucellosis. In the absence of a Gold Standard test, the performance of these diagnostic tests must be validated. The aim of the present study was to use the Bayesian latent class model to estimate the sensitivities and specificities of the AAT, 2-ME, FC and DBac diagnostic tests applied to blood samples and carcasses of animals suspected of bovine brucellosis, as well as the prevalence of the disease. The dataset used was obtained from the National Agricultural Laboratory of Minas Gerais (LANAGRO-MG). The tests were evaluated in two scenarios: individually and in combination. The models for the evaluation of the combined tests were adjusted considering the conditional independence between the four tests and also incorporating to the model the conditional dependence between the AAT, 2-ME and FC tests. Analyses were performed in R 3.2.5 using the *R2OpenBUGS* package. Regarding the evaluation of the combined tests, the results showed that the AAT, 2-ME and FC tests are conditionally independent. The FC test was the most sensitive, the DBac the least sensitive and the AAT, FC and DBac tests were the most specific. It was concluded that none of the four tests can be used alone for the diagnosis of bovine brucellosis. A low sensitivity was found for the AAT test, a result that diverges from the reports generally found in the literature. Therefore, it is recommended that continuous studies be carried out so that the decision-making of the researchers is not compromised. Additionally, it was concluded that the Bayesian latent class model employed allowed to estimate the parameters of interest satisfactorily.

1 INTRODUÇÃO

A brucelose bovina é uma doença que afeta vários países e no Brasil está presente em todas as regiões (BRASIL, 2006). Por ser o maior exportador de carne bovina (KIST et al., 2017) e o quinto maior produtor de leite do mundo (CONAB, 2017), o Brasil tem buscado combater os efeitos negativos causados por essa doença. Dentre os principais prejuízos econômicos provocados pela presença da brucelose bovina em um rebanho, destacam-se: abortamentos, nascimento de bezerros fracos (POESTER et al., 2010), redução na produção de leite e carne (BOTAHER NETO et al., 2009). Além disso tem-se a imposição de barreiras sanitárias e tarifárias ao comércio internacional de produtos de origem animal (SOLA et al., 2014).

A brucelose bovina é uma doença que apresenta sintomas semelhantes aos exibidos por animais supostamente infectados por outras doenças (BOTAHER NETO et al., 2009). Nesse caso, para realizar o diagnóstico da brucelose bovina, testes laboratoriais são necessários. Esses testes podem ser bacteriológicos ou sorológicos. Porém, os testes disponíveis para efetuar o diagnóstico da brucelose bovina não determinam corretamente a presença ou ausência da doença, uma vez que não são testes Padrão Ouro (POESTER et al., 2010), afetando assim a correta tomada de decisão dos profissionais envolvidos. Um teste diagnóstico é chamado Padrão Ouro se determina corretamente o verdadeiro estado de uma doença, isto é, se apresenta parâmetros de desempenho (sensibilidades e especificidades) iguais a 100% (SANOGO; ABATI; SAEGERMAN, 2014).

Com a implantação do Programa Nacional de Controle e Erradicação da Brucelose e da Tuberculose Animal (PNCEBT) pelo Ministério da Agricultura, Pecuária e Abastecimento (MAPA), que visa dentre outros, reduzir a prevalência da doença no território brasileiro (CLEMENTINO; AZEVEDO, 2016), ficou estabelecido, dentre outras ações, que o controle da doença se daria pela vacinação de fêmeas, diagnóstico e sacrifício dos animais identificados como portadores da doença (BRASIL, 2006). Como descrito anteriormente, não existe um teste Padrão Ouro para diagnosticar a brucelose bovina, portanto, o procedimento adotado é combinar dois ou mais testes, objetivando-se alcançar o diagnóstico definitivo (POESTER et al., 2010).

Para combater os efeitos da doença no território brasileiro, o PNCEBT estabeleceu os testes sorológicos: Antígeno Acidificado Tamponado (AAT), 2-Mercaptoetanol (2-ME) e Fixação de Complemento (FC), como testes oficiais para o diagnóstico da brucelose bovina. No entanto, testes bacteriológicos (DBac) também são utilizados (BRASIL, 2006). O AAT é utili-

zados por serem testes baratos e de fácil aplicação e, os testes 2-ME, FC e DBac, por apresentarem alta especificidade. Esses testes são conhecidos como diagnósticos imperfeitos, pois apresentam sensibilidade e/ou especificidade inferiores a 100% (SPEYBROECK et al., 2013), de modo que os parâmetros de desempenho desses testes necessitam de constante avaliação. Na ausência do teste Padrão Ouro, métodos estatísticos são utilizados para estimar os parâmetros de desempenho dos testes diagnósticos imperfeitos e neste contexto uma proposta é o modelo de classe latente com enfoque Bayesiano, sugerida por Joseph, Gyorkos e Coupal (1995). Nesse modelo, o verdadeiro estado da doença é uma variável denominada de latente, ou não observável, devido à ausência do teste Padrão Ouro.

Em modelos de classe latente, por suposição, os testes diagnósticos são condicionalmente independentes, porém, essa suposição é questionável na prática, principalmente se os testes combinados apresentam base biológica similar (GEORGIADIS et al., 2003), como os testes AAT, 2-ME e FC que detectam anticorpos contra a brucelose. Nesse caso, é necessário considerar o relaxamento da suposição de independência condicional, sendo que o teste DBac é considerado condicionalmente independente dos demais. Uma importante vantagem da abordagem Bayesiana frente à abordagem Frequentista é a questão da falta de identificabilidade (JONES et al., 2010), presente apenas na Frequentista. Por exemplo, o modelo de classe latente para um teste diagnóstico é não identificável, pois apresenta número de graus de liberdade menor que o número de parâmetros. No entanto, com o enfoque Bayesiano, todos os parâmetros são estimados, pois a abordagem tem a vantagem de incorporar informações a priori sobre a sensibilidade e a especificidade.

No estudo do desempenho de testes diagnósticos, quando há dúvidas sobre possível dependência condicional entre os testes, recomenda-se não ignorá-la, assim evita-se uma provável tendenciosidade nas estimativas das sensibilidades e especificidades (GEORGIADIS et al., 2003; DENDUKURI; BÉLISLE; JOSEPH, 2010). Nessa situação, o modelo de classe latente Bayesiano, considerando dependência condicional, tem sido uma ferramenta no desenvolvimento de estudos do desempenho de testes diagnósticos da brucelose bovina, como os realizados por: Matope et al. (2011), no Zimbábue; Sanogo et al. (2013), na Costa do Marfim; Getachew et al. (2016), na Etiópia; Praud et al. (2016), na França; entre outros. No entanto, no Brasil, não é de conhecimento, até o presente momento, que estudos para avaliar o desempenho de testes diagnósticos para brucelose bovina, considerando dependência condicional, tenham sido realizados.

Diante do exposto, neste trabalho, empregou-se o modelo de classe latente Bayesiano com o objetivo de estimar as sensibilidades e as especificidades dos testes diagnósticos AAT, 2-ME, FC e DBac, bem como a prevalência da doença, considerando-se que os testes tenham sido utilizados individualmente e combinados. Adicionalmente, na avaliação dos testes combinados, objetivou-se comparar o modelo de classe latente Bayesiano, considerando independência condicional entre os quatro testes com os modelos considerando dependência condicional entre os testes AAT x 2-ME, AAT x FC e 2-ME x FC, na ausência do Padrão Ouro. O ajuste do modelo de classe latente Bayesiano considerando-se dependência condicional entre os testes AAT, 2-ME e FC, foi realizado por meio da abordagem sugerida por Vacek (1985), que permite incorporar covariâncias entre os testes diagnósticos condicionadas ao verdadeiro estado da doença. Para avaliar o desempenho dos quatro testes diagnósticos, todos aplicados em amostras de sangue ou carcaça de animais suspeitos de brucelose bovina, o conjunto de dados utilizado foi obtido a partir dos materiais encaminhados pelo SIF (Serviço de Inspeção Federal) ao Laboratório Nacional Agropecuário de Minas Gerais (LANAGRO-MG).

2 REFERENCIAL TEÓRICO

2.1 Brucelose bovina

A Brucelose é uma doença infectocontagiosa causada por bactérias do gênero *Brucella* (BRASIL, 2006; MOL et al., 2012; AHASAN et al., 2017). Essa doença causa problemas sanitários, prejuízos econômicos na pecuária (BRASIL, 2006; SPICKLER, 2009b) e pode ser transmitida ao homem (GERESU; KASSA, 2016). A transmissão pode ocorrer diretamente (por meio do contato com animais infectados) ou indiretamente (por meio da ingestão de alimentos de origem animal, por exemplo, leite ou produtos lácteos não pasteurizados) (CORBEL, 2006; POESTER et al., 2010; GODFROID et al., 2011). Dentro do gênero *Brucella* pode-se apresentar as espécies: *Abortus*, em bovinos e bubalinos; *Melitensis* ou *Ovis*, em pequenos ruminantes; *Suis*, em suínos e *Canis*, em cães (BRASIL, 2006; SPICKLER, 2009b; GODFROID et al., 2011).

A espécie *Brucella Abortus* é mundialmente disseminada, com exceção de países como: Japão, Canadá, Austrália, Dinamarca, Nova Zelândia, Finlândia, Noruega, entre outros, onde houve erradicação (ROBINSON, 2003; SPICKLER, 2009a).

Sendo o Brasil, o país com o maior rebanho bovino comercial do mundo, com aproximadamente 215 milhões de cabeças (KIST et al., 2017), os prejuízos nos rebanhos de corte e também de leite causados pela brucelose bovina, doença predominante no território brasileiro (BRASIL, 2006; PAULA et al., 2015), é motivo de grandes preocupações. Dentre os principais prejuízos econômicos, tem-se: a redução na produção de leite, aborto, infertilidade, mortalidade perinatal, redução do peso, frequente substituição dos animais (SANOGO et al., 2013; ALVES et al., 2015; PAULA et al., 2015) e embargo às exportações de produtos de origem animal (SOLA et al., 2014), por exemplo, carne in natura (MALHEIROS, 2016). Outros prejuízos, incluem: a desvalorização dos animais de propriedades com brucelose, despesas com tratamentos e redução da produtividade de trabalhadores infectados (LAGE et al., 2008; PAULA et al., 2015).

Com o propósito de reduzir a prevalência e a incidência de novos casos de brucelose e tuberculose (CLEMENTINO; AZEVEDO, 2016), o Ministério da Agricultura, Pecuária e Abastecimento (MAPA) estabeleceu o Programa Nacional de Controle e Erradicação da Brucelose e da Tuberculose Animal (PNCEBT) em 2001 (BRASIL, 2006; POESTER et al., 2009). Dentre as recomendações do PNCEBT, destaca-se: a vacinação obrigatória de fêmeas com idade entre

três e oito meses (CLEMENTINO; AZEVEDO, 2016) e a certificação de propriedades livres ou monitoradas para a doença (SOLA et al., 2014). Os procedimentos para o reconhecimento de propriedades livres de brucelose segue às recomendações da Organização Mundial de Saúde Animal (OIE) (BRASIL, 2006).

2.1.1 Diagnóstico da brucelose bovina

Os principais sinais clínicos descritos em animais portadores da brucelose bovina são: aborto, nascimento de bezerros fracos e esterilidade de fêmeas e machos (LAGE et al., 2008; MANISH et al., 2013). Esses sinais são comuns em outras doenças bovinas, por exemplo: vibrose, leptospirose e listeriose (POESTER et al., 2010; ABUBAKAR; MANSOOR; ARSHED, 2012). Para diagnosticar a brucelose bovina é necessário a identificação do agente, que pode ser realizado por testes bacteriológicos (métodos diretos) ou pela detecção de anticorpos contra *B. abortus* por testes sorológicos (métodos indiretos) (BRASIL, 2006).

Os testes bacteriológicos incluem o isolamento e a identificação da *B. abortus* (BRASIL, 2006). A identificação é um procedimento lento, de elevados custos e há riscos de contaminação para os profissionais (LAGE et al., 2008; ABUBAKAR; MANSOOR; ARSHED, 2012), pois os materiais utilizados são provenientes de aborto (feto, conteúdo estomacal do feto e placenta) (LAGE et al., 2008; MOL et al., 2012) ou de secreções como leite, sêmen e líquido sinovial de articulações comprometidas (MOL et al., 2012; SOLA et al., 2014), por isso, poucos são os laboratórios que realizam esse teste (BRASIL, 2006).

Embora o teste bacteriológico apresente alta especificidade (MOL et al., 2012), esse teste é pouco sensível para diagnosticar a brucelose bovina (GERESU; KASSA, 2016). Como também não existe um teste sorológico capaz de identificar corretamente a doença, geralmente, o diagnóstico é realizado com base nos resultados de dois ou mais testes combinados (POESTER et al., 2010). Para realizar tal procedimento, habitualmente, utiliza-se um teste de triagem de fácil execução, barato e de boa sensibilidade (BRASIL, 2006). Se o teste de triagem classificar o animal como positivo, um teste confirmatório é realizado (POESTER et al., 2010). O teste confirmatório fornece boa sensibilidade, porém apresenta melhor especificidade quando comparado com o teste de triagem (BRASIL, 2006; POESTER et al., 2010).

O PNCEBT definiu os testes sorológicos: Antígeno Acidificado Tamponado (AAT), 2-Mercaptoetanol (2-ME) e Fixação de Complemento (FC), como oficiais para o diagnóstico da brucelose bovina (BRASIL, 2006), sendo o AAT como teste de triagem e os testes 2-ME

e FC como confirmatórios. No presente trabalho, por meio da colaboração do pesquisador Paulo Martins Soares Filho do LANAGRO-MG, foi realizado estudos sobre os parâmetros de desempenho (sensibilidade e especificidade) desses três testes diagnósticos, assim como do teste bacteriológico (DBac). Foram abordadas questões como a ausência do teste Padrão Ouro (seção 2.2) e possível dependência condicional entre os testes AAT, 2-ME e FC (seção 2.3.1). Portanto, espera-se que os resultados desse estudo possam contribuir com a correta tomada de decisão do pesquisador, especialmente, no sentido de evitar resultados falso negativos, o que possibilitaria a presença de animais doentes entre animais não doentes, algo que comprometeria o trabalho do programa de controle e erradicação da Brucelose (SOARES FILHO et al., 2011).

Os Laboratórios Nacionais Agropecuários (LANAGROS) são unidades oficiais do Ministério da Agricultura, Pecuária e Abastecimento (MAPA), que atuam em ações de monitoramento, controle e fiscalização de alimentos, bebidas e insumos produzidos e comercializados no Brasil (BRASIL, 2017). Dentre as unidades distribuídas em todas as regiões do país, o Laboratório Nacional Agropecuário de Minas Gerais (LANAGRO-MG), entre outras atividades: realiza diagnóstico das doenças dos animais; controla: produtos de origem animal, alimentos para animais, produtos veterinários e produtos de origem vegetal (BRASIL, 2011).

2.2 Medidas de desempenho de testes diagnósticos

2.2.1 Estimação com o teste Padrão Ouro

O desempenho dos testes diagnósticos é avaliado pela capacidade de identificar corretamente o estado da doença e é assunto de interesse em áreas como a medicina, medicina veterinária, fitopatologia, entre outras (HUI; ZHOU, 1998; GREINER; GARDNER, 2000b). Por exemplo, na medicina veterinária, testes com a capacidade de identificar corretamente o verdadeiro estado da doença podem contribuir em atividades de vigilância prescritas, na certificação de livre doença e para estimar a prevalência da doença (GREINER; GARDNER, 2000a). Na epidemiologia, tais testes detectam a presença da doença em estudos de fatores de risco e na avaliação de risco de transmissão da doença entre animais e seres humanos (GARDNER; HIETALA; BOYCE, 1996).

Geralmente, para avaliar o desempenho de um teste diagnóstico utiliza-se as medidas de sensibilidade e especificidade (MARTINEZ; ACHCAR; LOUZADA-NETO, 2006). Em tais estudos, outras medidas geralmente abordadas são as taxas de erros e a prevalência (HUI; ZHOU,

1998). Usualmente, essas medidas são obtidas por comparação com um teste de referência, denominado Padrão Ouro (ENØE; GEORGIADIS; JOHNSON, 2000), que é definido como um método de diagnóstico ou combinação de métodos que determina, sem erros, a presença ou ausência de uma determinada doença (GARDNER; HIETALA; BOYCE, 1996).

Para a definição da sensibilidade, especificidade e prevalência, considere a Tabela 1, elaborada conforme Joseph, Gyorkos e Coupal (1995). De acordo com Christensen et al. (2011), considere um teste diagnóstico que pode resultar em positivo ($T = 1$) ou negativo ($T = 0$), aplicado em uma amostra de n animais. Considere também a existência de um teste Padrão Ouro, em que o verdadeiro estado de cada animal, doente ($D = 1$) ou não doente ($D = 0$), é determinado pelo teste Padrão Ouro.

Tabela 1: Resultados de um teste diagnóstico considerando-se a existência do teste Padrão Ouro.

Teste	Padrão Ouro		Total
	$D = 1$	$D = 0$	
$T = 1$	a	b	$a + b$
$T = 0$	c	d	$c + d$
Total	$a + c$	$b + d$	n

Christensen et al. (2011) definem a , b , c e d , na Tabela 1, da seguinte maneira:

a = verdadeiros positivos - número de animais classificados corretamente como positivos;

b = falsos positivos - número de animais classificados erroneamente como positivos;

c = falsos negativos - número de animais classificados erroneamente como negativos;

d = verdadeiros negativos - número de animais classificados corretamente como negativos.

A sensibilidade e a especificidade dos testes diagnósticos em estudo serão denotadas seguindo o raciocínio de Martinez, Louzada-Neto e Pereira (2003). A prevalência da doença geralmente é designada pelas letras gregas θ e π , como verifica-se em Hui e Zhou (1998) e Dendukuri e Joseph (2001), respectivamente. A fim de evitar confusão de notação (ver seção 2.4), para este trabalho, a prevalência será indicada pela letra grega δ .

Com base nas informações apresentadas e com a definição de probabilidade condicional, a sensibilidade do teste diagnóstico, denotada por S , é dada por:

$$S = P(T = 1|D = 1) = \frac{P(T = 1, D = 1)}{P(D = 1)}, \quad (1)$$

isto é, a probabilidade condicional do teste fornecer resultado positivo, dado que o animal está doente (HUI; ZHOU, 1998). De acordo com a Tabela 1, a probabilidade S pode ser estimada como:

$$\widehat{S} = \frac{a}{a + c}.$$

A especificidade, denotada por E , é definida como:

$$E = P(T = 0|D = 0) = \frac{P(T = 0, D = 0)}{P(D = 0)}, \quad (2)$$

ou seja, é a probabilidade condicional do teste fornecer resultado negativo, dado que o animal não está doente (HUI; ZHOU, 1998). Essa probabilidade é estimada como:

$$\widehat{E} = \frac{d}{b + d}.$$

A prevalência é definida como a probabilidade de animais doentes na amostra (SANOGO; ABATIH; SAEGERMAN, 2014), sendo estimada da seguinte maneira:

$$\widehat{\delta} = P(D = 1) = \frac{a + c}{n}.$$

Em testes diagnósticos, dois tipos de erros de classificação são possíveis, denominadas de taxas de erros falso negativo e falso positivo (CHRISTENSEN et al., 2011). Esses erros podem ocorrer por várias razões: julgamento clínico subjetivo, imperfeições técnicas dos testes diagnósticos, adulteração deliberada, erros de interpretação por pesquisadores ou erros de escrita (WALTER; IRWING, 1988).

A taxa de erros falso negativo, denotada por $(1 - S)$ é definida como:

$$(1 - S) = P(T = 0|D = 1) = \frac{P(T = 0, D = 1)}{P(D = 1)}, \quad (3)$$

isto é, a probabilidade condicional do teste fornecer resultado negativo, dado que o animal está doente (HUI; ZHOU, 1998). Por meio da Tabela 1, o cálculo dessa probabilidade é determinado da seguinte forma:

$$(1 - \widehat{S}) = \frac{c}{a + c}.$$

A taxa de erros falso positivo, denotada por $(1 - E)$, é definida como:

$$(1 - E) = P(T = 1|D = 0) = \frac{P(T = 1, D = 0)}{P(D = 0)}, \quad (4)$$

é a probabilidade condicional do teste resultar positivo, dado que o animal não está doente (HUI; ZHOU, 1998). Sua estimativa é dada por:

$$(1 - \hat{E}) = \frac{b}{b + d}.$$

As taxas de erros falso negativo e falso positivo também são conhecidas, respectivamente, como complemento da sensibilidade e complemento da especificidade (TORRANCE-RYNARD; WALTER, 1997). As taxas de erros foram representadas seguindo o raciocínio de Collins e Huynh (2014).

2.2.2 Estimação na ausência do teste Padrão Ouro

Geralmente, testes diagnósticos Padrão Ouro são baseados em procedimentos de elevados custos, oferecem risco à integridade física ou até mesmo não existem (HUI; WALTER, 1980; GARDNER; HIETALA; BOYCE, 1996; MARTINEZ; LOUZADA-NETO; PEREIRA, 2003). Se tais testes diagnósticos não estão disponíveis, o verdadeiro estado da doença não pode ser diagnosticado, nesses casos, recomenda-se utilizar testes, denominados teste diagnóstico imperfeito, os quais são mais econômicos e práticos, (GARDNER; HIETALA; BOYCE, 1996; HUI; ZHOU, 1998). Um teste diagnóstico é imperfeito quando as medidas de sensibilidade e/ou especificidade são inferiores a 1 ou 100% (SPEYBROECK et al., 2013).

Na ausência do teste Padrão Ouro, não é possível determinar as estimativas da sensibilidade e especificidade diretamente (ENØE; GEORGIADIS; JOHNSON, 2000), conforme apresentado anteriormente nas expressões 1 e 2. Portanto, para estimar os parâmetros de desempenho de um teste diagnóstico, quando o teste Padrão Ouro não está disponível, é necessário utilizar métodos estatísticos inferenciais adequados. Dentro desse contexto, Hui e Walter (1980), empregaram o modelo de classe latente com enfoque Frequentista para estimar a sensibilidade e a especificidade, bem como a prevalência da doença. Os autores utilizaram dois testes para detectar a tuberculose em duas populações com distintas prevalências da doença. Basendo-se nos estudos de Hui e Walter (1980), Walter e Irwing (1988), desenvolveram o modelo para três testes diagnósticos e uma população. O termo “população”, utilizado pelos autores, na prática,

se refere as amostras ou subgrupos tomados da população de interesse.

Empregando o modelo de classe latente via enfoque Bayesiano, Joseph, Gyorkos e Coupal (1995) estimaram os parâmetros de desempenho na ausência do teste Padrão Ouro, no qual foram considerados um e dois testes diagnósticos e uma população. Utilizando os métodos propostos por Hui e Walter (1980) e Joseph, Gyorkos e Coupal (1995), com dois testes e uma população, Enøe, Georgiadis e Johnson (2000) verificaram que o método Bayesiano permite contornar limitações do método Frequentista, isto é, o enfoque Bayesiano possibilita, por meio das distribuições a priori, incorporar a opinião do especialista, algo que na análise Frequentista tende a ser omitida (BRANSCUM; GARDNER; JOHNSON, 2005). Do ponto de vista Frequentista, restrições são necessárias para estimar os parâmetros de interesse, por exemplo, no caso de um teste diagnóstico, para estimar a prevalência da doença, pressupõe-se que os valores da sensibilidade e especificidade são conhecidos, algo que na prática o pesquisador não tem disponível (SANOGO; ABATIH; SAEGERMAN, 2014). Outra vantagem do enfoque Bayesiano é a de que todos os parâmetros de interesse podem ser estimados, uma vez que para cada um deles é atribuído uma distribuição à priori, sendo assim a necessidade de restrições impostas no enfoque Frequentista é eliminada, sendo que essas restrições frequentemente são realizadas de forma arbitrária (JOSEPH; GYORKOS; COUPAL, 1995).

O desenvolvimento desses trabalhos mostrou a aplicabilidade dos métodos estatísticos na validação dos testes diagnósticos, na ausência do teste Padrão Ouro, permitindo diversos estudos para diagnosticar infecções e doenças, como: paratuberculose bovina (GARDNER et al., 2000); pneumonia por mycoplasma em suínos (HANSON; JOHNSON; GARDNER, 2003); clamídia (DENDUKURI; HADGU; WANG, 2009); infecção por *Strongyloides stercoralis* e transmissão de tuberculose (DENDUKURI; BÉLISLE; JOSEPH, 2010); HIV - vírus da imunodeficiência humana (BROWN, 2010); brucelose suína (PRAUD et al., 2012) e bovina (SANOGO et al., 2013); dengue (PAN-NGUM et al., 2013); brucelose em búfalo africano (GORSICH et al., 2015); obesidade (O'NEILL, 2015), entre outras.

Nessa seção verificou-se como determinar as medidas de diagnóstico na presença do teste Padrão Ouro e também foram citadas metodologias estatísticas que podem ser empregadas quando não se dispõe do teste Padrão Ouro. A seguir será apresentado o modelo de classe latente.

2.3 Modelo de classe latente

Em muitas áreas de pesquisas científicas pode ser de interesse estudar variáveis que não podem ser observadas diretamente, tais variáveis são denominadas latentes (McCUTCHEON, 1987; VERMUNT; MAGIDSON, 2004). Uma vez que não podem ser observadas diretamente, as variáveis latentes são obtidas indiretamente por meio de duas ou mais variáveis observáveis (NEELY-BARNES, 2010; COLLINS; LANZA, 2010). As variáveis latentes e observadas podem ser quantitativas ou qualitativas (categóricas). Quando as variáveis latentes e observadas são categóricas, emprega-se o modelo de classe latente (VERMUNT; MAGIDSON, 2004; SKRONDAL; RABE-HESKETH, 2007).

Essencialmente, o modelo de classe latente é utilizado com o objetivo de identificar um conjunto de classes latentes mutuamente exclusivos, com base em uma série de variáveis observadas (McCUTCHEON, 1987; LANZA et al., 2007; REITSMA et al., 2009; VIDAL et al., 2012), de modo que, as relações existentes entre essas variáveis observadas podem ser explicadas pelas variáveis latentes (YOUNG, 1982; SMEDEN et al., 2014).

O uso de modelos de classe latente para avaliar o desempenho de testes diagnósticos, quando o teste Padrão Ouro não está disponível, tem sido constante em áreas como medicina e medicina veterinária (SMEDEN et al., 2014). Após a publicação do trabalho desenvolvido por Hui e Walter (1980), a maioria das publicações posteriores aplicaram o modelo proposto pelos autores a dados com variados números de testes (COLLINS; HUYNH, 2014). Como exemplos: Hadgu e Qu (1998) que estimaram a sensibilidade e especificidade de seis testes de triagem para detecção de *Chlamydia trachomatis*, bactéria causadora da Clamídia, DST que atinge homens e mulheres; Albert e Dodd (2004) avaliaram cinco testes para verificar a presença de cáries incipientes; Toft, Jørgensen e Højsgaard (2005) aplicaram o modelo de Hui e Walter em dados de paratuberculose bovina utilizando dois testes; Menten, Boelaert e Lesaffre (2008) avaliaram quatro testes diagnósticos para a leishmaniose visceral, entre outros.

Em modelos de classe latente os testes diagnósticos são indicadores imperfeitos do verdadeiro estado da doença que é desconhecido (RINDSKOPF; RINDSKOPF, 1986; YANG; BECKER, 1997). Os resultados observados em diferentes testes imperfeitos são influenciados pela mesma variável latente e quanto maior o número de testes, maior será o conhecimento acerca do verdadeiro estado da doença (HADGU; DENDUKURI; HILDEN, 2005). Na literatura, esse modelo tem sido parametrizado de duas formas: em termos de probabilidade e em termos de modelo log-linear (COLLINS; LANZA, 2010). A parametrização adotada neste tra-

balho será em termos de probabilidade e os parâmetros a serem estimados, conforme Collins e Lanza (2010), são: a probabilidade (não condicional) de um animal pertencer a uma certa classe latente (doente ou não doente) e a probabilidade condicional do teste resultar positivo ou negativo dado que o animal está doente ou não doente.

Para especificar o modelo de classe latente, segundo Menten, Boelaert e Lesaffre (2008), sendo $i = 1, \dots, n$ animais e $j = 1, \dots, J$ testes diagnósticos, considere T_{ij} a variável que indica o resultado do j -ésimo teste aplicado no i -ésimo animal, tal que, o teste pode resultar em negativo ($t_{ij} = 0$) ou positivo ($t_{ij} = 1$). Equivalentemente, seja D_i o verdadeiro, porém, desconhecido estado da doença do i -ésimo animal, sendo ($d_i = 0$), se animal for não doente ou ($d_i = 1$), se animal for doente. Assim, a distribuição conjunta dos resultados dos J testes diagnósticos é:

$$\begin{aligned} P(T_{i1} = t_{i1}, \dots, T_{iJ} = t_{iJ}) &= P(T_{i1} = t_{i1}, \dots, T_{iJ} = t_{iJ}, D_i = 0) + P(T_{i1} = t_{i1}, \dots, T_{iJ} = t_{iJ}, D_i = 1) \\ &= \sum_{d_i=0,1} P(T_{i1} = t_{i1}, \dots, T_{iJ} = t_{iJ}, D_i = d_i) \\ &= \sum_{d_i=0,1} P(D_i = d_i)P(T_{i1} = t_{i1}, \dots, T_{iJ} = t_{iJ}|D_i = d_i). \end{aligned} \quad (5)$$

Sob a pressuposição de independência condicional entre os testes diagnósticos, dado o verdadeiro estado da doença, tem-se que:

$$P(T_{i1} = t_{i1}, \dots, T_{iJ} = t_{iJ}|D_i = d_i) = \prod_{j=1}^J P(T_{ij} = t_{ij}|D_i = d_i). \quad (6)$$

Combinando as expressões (5) e (6), obtém-se:

$$P(T_{i1} = t_{i1}, \dots, T_{iJ} = t_{iJ}) = \sum_{d_i=0,1} P(D_i = d_i) \prod_{j=1}^J P(T_{ij} = t_{ij}|D_i = d_i). \quad (7)$$

Como já mencionado, cada teste fornecerá apenas dois resultados, positivo ou negativo, portanto admite-se que a variável tem distribuição Bernoulli, isto é:

$$P(T_{ij} = t_{ij}) = p^{t_{ij}}(1 - p)^{1-t_{ij}}, \quad (8)$$

em que $p = P(T_{ij} = 1)$ e $(1 - p) = P(T_{ij} = 0)$. Dado o verdadeiro estado da doença no animal i , a probabilidade de sucesso é dada por $p_{jd_i} = P(T_{ij} = 1|D_i = d_i)$, caso contrário, $(1 - p_{jd_i}) = P(T_{ij} = 0|D_i = d_i) = 1 - P(T_{ij} = 1|D_i = d_i)$. Também já foi referido que as classes (conjuntos ou grupos) latentes são mutuamente exclusivas e exaustivas, em outras palavras, cada animal pertence a uma única classe latente, isso significa que $\sum_{d_i} P(D_i = d_i) = 1$ (LANGEHEINE, 1988; COLLINS; LANZA, 2010). Seja $P(D_i = d_i) = \delta_{d_i}$ e sabendo que cada animal é representado por uma distribuição Bernoulli, a expressão (7) pode ser reescrita

como:

$$P(T_{i1} = t_{i1}, \dots, T_{iJ} = t_{iJ}) = \sum_{d_i=0,1} \delta_{d_i} \prod_{j=1}^J p_{jd_i}^{t_{ij}} (1 - p_{jd_i})^{1-t_{ij}}, \quad (9)$$

sendo, $\delta_{d=1} = \delta$ a prevalência da doença, $\delta_{d=0} = (1 - \delta)$ a probabilidade de animais não doentes na amostra e t_{ij} conforme já mencionado. Segundo Yang e Becker (1997), o modelo da expressão (9) é adequado para estimar os parâmetros de desempenho de testes diagnósticos, uma vez que o modelo é parametrizado de acordo com as probabilidades que definem a sensibilidades, $S_j = P(T_{ij} = 1 | D_i = 1)$; especificidades, $E_j = P(T_{ij} = 0 | D_i = 0)$ e a prevalência da doença, $\delta = P(D_i = 1)$. Assim, a expressão (9) pode ser reescrita da seguinte forma:

$$P(T_{i1} = t_{i1}, \dots, T_{iJ} = t_{iJ}) = \delta \prod_{j=1}^J S_j^{t_{ij}} (1 - S_j)^{1-t_{ij}} + (1 - \delta) \prod_{j=1}^J E_j^{1-t_{ij}} (1 - E_j)^{t_{ij}}. \quad (10)$$

Embora o modelo de classe latente seja uma ferramenta para estimar os parâmetros de desempenho de testes diagnósticos em diferentes doenças e números de testes, dois problemas devem ser considerados: o questionamento da suposição de independência condicional entre os testes diagnósticos e a identificabilidade do modelo. Estes dois conceitos serão abordados a seguir.

2.3.1 Independência condicional

A independência condicional é uma pressuposição fundamental para modelos de classe latente, pois determina que, condicionada à variável latente, que informa o verdadeiro estado da doença, as variáveis observadas (os resultados dos testes) são independentes (COLLINS; LANZA, 2010). Dois testes diagnósticos são condicionalmente independentes, se a probabilidade do resultado do segundo teste (T_{i2}) não depende dos resultados do primeiro teste (T_{i1}), isto é, se o primeiro teste resulta em positivo ou negativo entre os animais doentes ou não doentes (GARDNER et al., 2000; TOFT; JØRGENSEN; HØJSGAARD, 2005; SANOGO; ABATIH; SAEGERMAN, 2014). Portanto, dois testes diagnósticos são condicionalmente independentes se:

$$\begin{aligned} P(T_{i1}, T_{i2} | D) &= \frac{P(T_{i1}, T_{i2}, D)}{P(D)} \\ &= \frac{P(\cancel{D}) P(T_{i1} | D) P(T_{i2} | T_{i1}, D)}{P(\cancel{D})} \\ &= P(T_{i1} | D) P(T_{i2} | T_{i1}, D). \end{aligned} \quad (11)$$

Sob independência condicional, $P(T_{i2} | T_{i1}, D) = P(T_{i2} | D)$, então:

$$P(T_{i1}, T_{i2}|D) = P(T_{i1}|D)P(T_{i2}|D). \quad (12)$$

Porém, geralmente a suposição de independência condicional é questionável na prática (VACEK, 1985; YANG; BECKER, 1997; HUI; ZHOU, 1998; MENTEN; BOELAERT; LE-SAFFRE, 2008), especialmente se os testes apresentam base biológica similar (VACEK, 1985; GARDNER et al., 2000; DENDUKURI; BÉLISLE; JOSEPH, 2010). Por exemplo, os testes diagnósticos em estudo: AAT, 2-ME e FC podem ser condicionalmente dependentes, pois são baseados na detecção de anticorpos contra a *Brucella abortus*, algo que pode inibir uma reação ou provocar uma falsa reação em um dos testes (VACEK, 1985). Sendo assim, considerar independência condicional entre os testes diagnósticos, quando não existe, pode resultar em estimativas tendenciosas para a sensibilidade e especificidade (GEORGIADIS et al., 2003; SMEDEN et al., 2014). Nesse sentido, alguns modelos foram propostos para o relaxamento da suposição de independência condicional, permitindo que a dependência condicional entre os resultados dos testes diagnósticos, na ausência do Padrão Ouro, possa ser incorporada no modelo.

Um dos primeiros trabalhos com essa abordagem foi desenvolvido por Vacek (1985), que baseado no modelo de Hui e Walter (1980), avaliou o efeito da dependência condicional entre dois testes por meio de dois parâmetros: a covariância entre os testes dado que o animal é portador da doença e a covariância entre os testes dado que o animal não é portador da doença. Por meio de simulação e seguindo a metodologia proposta por Vacek, Torrance-Rynard e Walter (1997), verificaram que as especificidades e especialmente as sensibilidades de dois testes condicionalmente dependentes são superestimadas, o que levará a falsas conclusões em relação ao desempenho dos testes. Na literatura, modelos em que a dependência é construída usando covariâncias são conhecidos como modelos de efeito fixo.

Utilizando outra abordagem, Qu, Tan e Kutner (1996), propuseram o modelo de efeitos aleatórios para avaliar quatro testes diagnósticos. Essa formulação consiste em modelar a dependência condicional pela inclusão de uma variável latente contínua com distribuição normal padrão. Também empregando quatro testes diagnósticos, Yang e Becker (1997), usaram o modelo log-linear para explicar a dependência. Aplicando o enfoque Bayesiano, Dendukuri e Joseph (2001), realizaram inferências sobre a prevalência da doença, sensibilidade e especificidade em dois testes diagnósticos, supondo dependência condicional, por meio de modelos de efeito fixo e aleatório.

2.3.2 Identificabilidade

Um problema geralmente encontrado na análise de classe latente é a identificabilidade do modelo (DENDUKURI; HADGU; WANG, 2009). Um modelo é estritamente (ou globalmente) identificável se diferentes estimativas dos parâmetros correspondem a diferentes distribuições de probabilidade (WIERINGEN, 2005; WANG; ZHOU, 2014), isto é, um modelo F é identificável se, para quaisquer dois valores dos parâmetros $\theta_1, \theta_2 \in \Theta$, tal que $\theta_1 \neq \theta_2$, então $F(\theta_1) \neq F(\theta_2)$, sendo Θ o espaço paramétrico. E se distintos valores dos parâmetros resultam na mesma distribuição de probabilidade, em outras palavras, se existem $\theta_1 \neq \theta_2$ tal que $F(\theta_1) = F(\theta_2)$, então o modelo é dito não identificável (JONES et al., 2010; WANG; ZHOU, 2014). A implicação prática da não identificabilidade é a possibilidade de distintos valores de parâmetros resultarem no mesmo valor da função de verossimilhança da amostra e consequentemente, inviabilizar uma adequada estimação dos parâmetros.

Uma condição necessária para a identificabilidade do modelo é que o número de graus de liberdade (gl) seja maior ou igual ao número de parâmetros (r) (COLLINS; HUYNH, 2014).

Para maior entendimento, considere o exemplo dado por Huang (2005), em que deseja-se classificar pessoas em três estados não observáveis: $D = 1 =$ não depressiva; $D = 2 =$ levemente depressiva e $D = 3 =$ severamente depressiva. Três testes psicológicos ($J = 3$) foram utilizados para diagnosticar a doença e os resultados são dados por: $T_j = 1 =$ positivo ou $T_j = 0 =$ negativo. Assim, a distribuição conjunta dos resultados dos testes é:

$$P(T_1 = t_1, \dots, T_3 = t_3) = \sum_{d=1}^3 P(D = d) \prod_{j=1}^3 P(T_j = 1|D = d)^{t_j} P(T_j = 0|D = d)^{1-t_j}.$$

Segundo o autor, este modelo tem $2^J - 1 = 7$ graus de liberdade e 11 parâmetros a serem estimados, isto é, $P(D = 2)$, $P(D = 3)$ e uma $P(T_j = 1|D = d)$ para cada estado de depressão d e teste j . Sendo $gl < r$, o modelo não possui condição básica para identificabilidade. Neste caso, restrições sobre os parâmetros são necessárias e as consideradas foi supor que $P(T_j = 1|D = 3) = P(T_j = 1|D = 2) = P(T_j = 1|D = 1); j = 2, 3$. Com isso, apenas cinco parâmetros serão estimados numa condição de modelo identificável.

No entanto, modelos de classe latente não são estritamente identificáveis (JONES et al., 2010), e sim centrados em uma forma mais fraca de identificabilidade, denominada “identificabilidade local” (McHUGH, 1956; GOODMAN, 1974; WANG; ZHOU, 2014; XU; ZHANG, 2016). Um modelo F é localmente identificável em um parâmetro $\theta_0 \in \Theta$ se existe alguma vizinhança de θ_0 , N_{θ_0} , tal que para qualquer $\theta \in N_{\theta_0}$, $F(\theta) \neq F(\theta_0)$ (WANG; ZHOU, 2014), ou

melhor, dentro de uma vizinhança suficientemente pequena no espaço paramétrico, a identificabilidade local requer que o parâmetro seja único (ALLMAN; MATIAS; RHODES, 2009). Dentre os primeiros trabalhos com essa abordagem, tem-se os desenvolvidos por McHugh (1956), que propôs condições para a identificabilidade local em modelos com variáveis observadas e latentes dicotômicas e Goodman (1974), que generalizou essa proposta para modelos com variáveis politômicas. A proposta sugerida considera que se o posto da matriz Jacobiana for igual ao número de colunas, isto é, o número de parâmetros a serem estimados, então o modelo é localmente identificável e esta é uma condição suficiente (GOODMAN, 1974). Porém, essa proposta não permite afirmar que a condição é satisfeita apenas empiricamente, isto é, apenas sobre os dados em estudo (DENDUKURI; HADGU; WANG, 2009).

O problema da falta de identificabilidade é uma questão presente no enfoque Frequentista e Bayesiano. Do ponto de vista Frequentista, a solução pode ser a partir de restrições sobre os parâmetros e assim o espaço paramétrico é reduzido (XIE; CARLIN, 2004), como mostrado no exemplo anteriormente, enquanto que no enfoque Bayesiano, a solução será a partir da utilização de distribuições a priori para os parâmetros de interesse (JOSEPH; GYORKOS; COUPAL, 1995).

Embora a condição $gl \geq r$ seja necessária para a identificabilidade do modelo, vários autores relatam que não é suficiente, mesmo quando satisfeita (McHUGH, 1956; GOODMAN, 1974; BASSI, 1997; FORMANN, 2006; ALLMAN; MATIAS; RHODES, 2009; JONES et al., 2010; WANG; ZHOU, 2014), porém é amplamente aceita na literatura, pois apesar de algumas tentativas para estabelecer critérios gerais para investigar a identificabilidade local em modelos de classe latente, ainda não existe consenso à respeito. Para o presente trabalho, a condição $gl \geq r$ será adotada para indicar a identificabilidade dos modelos. Nos casos em que não for satisfeita será considerado que a informação a priori, fornecida com base na experiência do pesquisador do LANAGRO-MG, é suficiente para realizar inferências a partir dos modelos não identificáveis em estudo, conforme sugere Jones et al. (2010).

2.4 Enfoque Bayesiano

Para estimar os parâmetros de desempenho de testes diagnósticos e a prevalência da doença utilizando modelos de classe latente, tanto o enfoque Frequentista quanto o Bayesiano são empregados, no entanto, com o passar dos anos, o enfoque Bayesiano vem sendo largamente aplicado na área de validação de testes diagnósticos (GARDNER, 2002). A empregabilidade

do enfoque Bayesiano para avaliar o desempenho de testes diagnósticos é, particularmente, pela possibilidade de informações a priori sobre a prevalência da doença, sensibilidade e especificidade dos testes serem incorporadas (BRANSCUM; GARDNER; JOHNSON, 2005). Essas informações podem ser definidas com base na opinião de especialistas ou resultados de pesquisas anteriores (SPEYBROECK et al., 2013).

Na metodologia Bayesiana os dados \mathbf{X} são considerados fixos e os parâmetros θ são variáveis aleatórias (SANOGO; ABATIH; SAEGERMAN, 2014). Toda a incerteza inicial sobre os parâmetros é quantificada por meio de probabilidades (GELMAN et al., 2004), e posteriormente, qualquer inferência à respeito da quantidade desconhecida θ , será com base em uma distribuição a posteriori para θ obtida por meio do teorema de Bayes, descrito da seguinte maneira:

$$\pi(\theta|\mathbf{x}) = \frac{\pi(\theta)f(\mathbf{x}|\theta)}{\int \pi(\theta)f(\mathbf{x}|\theta)d\theta}, \quad (13)$$

em que, $\pi(\theta)$ representa as distribuições a priori do vetor de parâmetros θ ; $f(\mathbf{x}|\theta)$ a função densidade no caso contínuo ou distribuição de probabilidade no caso discreto de \mathbf{X} , sendo interpretada como a distribuição condicional de \mathbf{X} dado θ e $\pi(\theta|\mathbf{x})$ é a distribuição condicional de θ dado $\mathbf{X} = \mathbf{x}$, denominada distribuição a posteriori.

Para um valor fixo de \mathbf{x} , a função $f(\mathbf{x}|\theta)$, conhecida como função de verossimilhança, fornece a verossimilhança para os possíveis valores de θ (GELMAN et al., 2004; CHRISTENSEN et al., 2011). Na equação (13), o denominador pode ser omitido, pois não depende de θ e como \mathbf{x} é fixo, pode ser considerado como uma constante normalizadora da distribuição a posteriori, que resulta na equação (14), isto é,

$$\pi(\theta|\mathbf{x}) \propto \pi(\theta)f(\mathbf{x}|\theta). \quad (14)$$

Portanto, a metodologia Bayesiana para realizar inferência sobre os parâmetros θ , combina informações a priori $\pi(\theta)$ com os dados $f(\mathbf{x}|\theta)$ para obter a distribuição a posteriori $\pi(\theta|\mathbf{x})$ de θ (LESAFFRE; SPEYBROECK; BERKVENNS, 2007).

2.4.1 Função de verossimilhança

A função de verossimilhança, $f(\mathbf{x}|\theta)$, é um componente importante no teorema de Bayes, pois é por meio desta função que os dados \mathbf{x} modificam o conhecimento a priori de

θ (BOX; TIAO, 1973). A função de verossimilhança é dada por:

Definição 2.4.1. Seja $\mathbf{x} = (x_1, \dots, x_n)$ uma amostra aleatória observada de uma distribuição que depende de um ou mais parâmetros desconhecidos θ , com função densidade de probabilidade (fdp) ou função de probabilidade (fp), denotada por $f(\mathbf{x}|\theta)$. Então, a fdp ou fp conjunta de \mathbf{x} , definida por:

$$\begin{aligned} L(\theta|\mathbf{x}) &= f(x_1|\theta) \times \dots \times f(x_n|\theta) \\ &= \prod_{i=1}^n f(x_i|\theta), \end{aligned} \quad (15)$$

é chamada de função de verossimilhança (CASELLA; BERGER, 2010).

2.4.2 Distribuição a priori

A distribuição a priori representa uma situação em que as informações acerca dos parâmetros de interesse são obtidas antes da realização do experimento, isto é, a distribuição a priori expressa todo o conhecimento sobre θ antes da observação dos dados amostrais \mathbf{x} (BOX; TIAO, 1973; PRESS, 2003; GELMAN et al., 2004; BOLSTAD, 2010; CHRISTENSEN et al., 2011).

A escolha da distribuição a priori é talvez o ponto mais crítico e criticado na Inferência Bayesiana, pois a sua construção depende da natureza do conhecimento (informativo ou não informativo) que se baseia o especialista e/ou estatístico, ao passo que quantificar essa informação probabilisticamente, a ser incorporada ao teorema de Bayes, tem sido algo discutido extensivamente na literatura (PAULINO; TURKMAN; MURTEIRA, 2003; PRESS, 2003; ROBERT, 2007). Nesse sentido, algumas formas de especificação da distribuição a priori são frequentemente abordadas na literatura, dentre elas, as prioris conjugadas. A ideia de prioris conjugadas é que as distribuições a priori e posteriori pertençam a mesma família de distribuição (PRESS, 2003), dessa forma, a atualização das informações sobre os parâmetros do vetor θ implica unicamente em uma mudança nos hiperparâmetros (GELMAN et al., 2004; LEE, 2007; CHRISTENSEN et al., 2011). Os hiperparâmetros são os parâmetros da distribuição a priori, essa denominação é utilizada para diferenciá-los dos parâmetros do vetor θ (ROBERT, 2007).

Geralmente, em estudos para avaliar o desempenho de testes diagnósticos, adota-se a densidade Beta como distribuição a priori para os parâmetros, pois a utilização dessa distribuição simplifica os cálculos e também é bastante flexível em acomodar diversas formas (ENØE; GEORGIADIS; JOHNSON, 2000; ARAGON; MARTINEZ; ACHCAR, 2010), além de ser

uma distribuição conjugada da distribuição a posteriori. Assim, uma variável aleatória terá distribuição beta se satisfazer a definição a seguir.

Definição 2.4.2. Sejam $\alpha, \beta > 0$ e θ uma variável aleatória com função densidade de probabilidade dada por:

$$f(\theta) = \begin{cases} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1} (1 - \theta)^{\beta-1}, & 0 < \theta < 1 \\ 0, & \text{caso contrário,} \end{cases} \quad (16)$$

então θ tem distribuição beta com parâmetros α e β . Sendo $\Gamma(\cdot)$, a função Gama.

Sendo o vetor de parâmetros θ , tem-se que $\text{Beta}(\alpha_\theta, \beta_\theta)$ representa a distribuição beta a priori de cada parâmetro de interesse do vetor, em que α_θ e β_θ são os hiperparâmetros. Uma importante vantagem ao utilizar a densidade Beta como distribuição a priori para modelar porcentagens, está no fato de sua região de densidade, também denominada suporte de $f(\theta)$, estar definida em $(0, 1)$, o que corresponde ao intervalo dos parâmetros de interesse nesse estudo (JOSEPH; GYORKOS; COUPAL, 1995).

Quando dois testes diagnósticos são condicionalmente dependentes, dois parâmetros são incorporados na análise, as covariâncias entre os testes dado a presença $covs_{12}$ e ausência $cove_{12}$ da doença. De acordo com Dendukuri e Joseph (2001), esses parâmetros têm distribuição beta generalizada a priori, $covs_{12} \sim \text{genbeta}(\alpha_{covs_{12}}, \beta_{covs_{12}})$ e $cove_{12} \sim \text{genbeta}(\alpha_{cove_{12}}, \beta_{cove_{12}})$. A parametrização das covariâncias condicionais está conforme Vacek (1985).

Definição 2.4.3. Uma variável aleatória θ com função de densidade de probabilidade:

$$f(\theta) = \begin{cases} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{(\theta - L_1)^{\alpha-1} (L_2 - \theta)^{\beta-1}}{(L_2 - L_1)^{\alpha+\beta-1}}, & \text{se } L_1 \leq \theta \leq L_2 \\ 0, & \text{caso contrário,} \end{cases} \quad (17)$$

é dita ter distribuição beta generalizada definida no intervalo $[L_1, L_2]$, com $\alpha, \beta > 0$. Sendo $\Gamma(\cdot)$, a função Gama.

A distribuição beta generalizada é uma densidade beta redimensionada em um intervalo diferente de $[0, 1]$ (DENDUKURI; JOSEPH, 2001; BRANSCUM; GARDNER; JOHNSON, 2005).

Geralmente, não existe informação a priori para os parâmetros $covs_{12}$ e $cove_{12}$. Nesse caso, uma solução é utilizar a distribuição uniforme a priori para as duas covariâncias, tal que, os limites do intervalo da distribuição uniforme do parâmetro $covs_{12}$ e $cove_{12}$ são descritos

em função da sensibilidade e especificidade, respectivamente (DENDUKURI; JOSEPH, 2001; BRANSCUM; GARDNER; JOHNSON, 2005). Neste trabalho, devido à ausência de informações a priori sobre as covariâncias, adotou-se a estratégia citada anteriormente. Para mais detalhes, seção 3.3.

2.4.3 Distribuição a posteriori

A distribuição a posteriori reflete a atualização do conhecimento a priori sobre θ após observar os dados amostrais \mathbf{x} (BERGER, 1985; CHRISTENSEN et al., 2011), de modo que essa distribuição contém todas as informações sobre θ , e essas informações podem ser resumidas por meio de valores numéricos como: média, mediana, moda (GELMAN; RUBIN, 1992; GELMAN et al., 2004) ou por intervalos de credibilidade (BRANSCUM; GARDNER; JOHNSON, 2005).

Sendo o objetivo realizar inferências sobre o vetor de parâmetros $\theta = (\theta_1, \dots, \theta_k)$, é necessário obter as distribuições marginais a posteriori de cada parâmetro do vetor θ . Por exemplo, para o parâmetro θ_1 e considerando a expressão (14), tem-se que:

$$\pi(\theta_1|\mathbf{x}) = \int \pi(\theta|\mathbf{x})d\theta_2 \dots d\theta_k, \quad (18)$$

é a distribuição marginal a posteriori do parâmetro θ_1 , que é obtida integrando a distribuição conjunta a posteriori em relação aos parâmetros do vetor, exceto θ_1 (GELMAN et al., 2004).

No entanto, frequentemente, a distribuição conjunta a posteriori é analiticamente intratável, pois envolve a resolução de integrais complexas, especialmente quando existe um considerável número de parâmetros (BOLSTAD, 2010; PEREIRA et al., 2012) ou a mesma não pertence a uma família de distribuições conhecidas, tal como a normal, beta ou gama (BRANSCUM; GARDNER; JOHNSON, 2005). Nesses casos, métodos de simulação são necessários para resumir informações da distribuição conjunta a posteriori. Na inferência Bayesiana esse processo pode ser realizado utilizando-se os métodos de Monte Carlo via Cadeias de Markov (MCMC) (GELMAN et al., 2004; ROBERT; CASELLA, 2010). Esses métodos são fundamentados na simulação de cadeias de Markov que converge para a densidade a posteriori de interesse e, os algoritmos Metropolis-Hastings e Amostrador de Gibbs são comumente empregados para este fim (PRESS, 2003). Dentre os algoritmos citados, o Amostrador de Gibbs será abordado a seguir, pois a distribuição beta é conjugada à distribuição binomial conforme declara Christensen et al. (2011).

para monitorar a convergência, por exemplo, os propostos por Gelman e Rubin (1992) e Raftery e Lewis (1992).

Antes de apresentarmos uma breve introdução sobre os dois critérios citados anteriormente, algumas especificações para a execução dos algoritmos são necessárias, por exemplo, o número de iterações para o processo de simulação ser executado (COWLES; CARLIN, 1996), o descarte das primeiras iterações na simulação, com o objetivo de evitar a influência dos valores iniciais, conhecido como período de aquecimento, *burn-in* (BROOKS, 1998; GELMAN et al., 2004) e o espaçamento entre as observações, *thin*, afim de obter amostras independentes (GELMAN et al., 2004).

2.4.5.1 Critério de Gelman e Rubin

Proposto por Gelman e Rubin (1992), a ideia do diagnóstico é simular m cadeias paralelas, cada uma com $2n$ iterações e com diferentes valores iniciais (CARLIN; LOUIS, 2008). O método baseia-se na comparação das variâncias entre e dentro das cadeias, que é conhecido como fator de redução de escala potencial (\hat{R}) (KAPLAN, 2014). Sendo as cadeias executadas de modo que as n primeiras iterações são descartadas, para evitar o *burn-in* (BROOKS; ROBERTS, 1998), considere a variável aleatória θ , que possui média μ e variância σ^2 sob a distribuição de interesse, e suponha que existe um estimador não tendencioso $\hat{\mu}$ para μ . Seja θ_{tj} , a t -ésima observação dentre as n iterações de θ na cadeia j , e também $\hat{\mu} = \bar{\theta}_{..}$ (BROOKS; GELMAN, 1998), então, as variâncias entre (B) e dentro (W) das cadeias são obtidas, respectivamente, por:

$$B = \frac{n}{m-1} \sum_{j=1}^m (\bar{\theta}_{.j} - \bar{\theta}_{..})^2 \quad \text{e} \quad W = \frac{1}{m(n-1)} \sum_{j=1}^m \sum_{t=1}^n (\theta_{tj} - \bar{\theta}_{.j})^2,$$

em que $\bar{\theta}_{.j} = \frac{1}{n} \sum_{t=1}^n \theta_{tj}$ é a média das observações da cadeia j e $\bar{\theta}_{..} = \frac{1}{m} \sum_{j=1}^m \bar{\theta}_{.j}$ a média geral (GELMAN et al., 2004; GAMERMAN; LOPES, 2006). Obtidas as estimativas acima, a variância de θ , é dada pela a média ponderada de B e W , isto é:

$$\hat{\sigma}_{\theta}^2 = \left(\frac{n-1}{n} \right) W + \left(\frac{1}{n} \right) B.$$

Assim, conforme Gelman et al. (2004) e Kaplan (2014), um indicador de convergência pode ser dado pelo estimador de fator de redução de escala potencial, que é obtido pela expressão a seguir:

$$\hat{R} = \sqrt{\frac{\hat{\sigma}_\theta^2}{W}}.$$

Se \hat{R} é próximo de 1, pode-se concluir que cada m conjuntos de n observações simuladas aproxima-se da distribuição a posteriori de interesse (GELMAN; RUBIN, 1992; BROOKS; ROBERTS, 1998; BROOKS; GELMAN, 1998).

2.4.5.2 Critério de Raftery e Lewis

Proposto por Raftery e Lewis (1992), o critério fornece o número de iterações necessários para alcançar a convergência (PAULINO; TURKMAN; MURTEIRA, 2003). Ademais, o método sugere o número de iterações para o período de aquecimento e o valor do espaçamento entre as observações (GAMERMAN; LOPES, 2006). Outra importante informação fornecida pelo método é o fator de dependência (F_D), o qual mede o aumento no número de iterações necessárias para alcançar a convergência (BEST; COWLES; VINES, 1995). Para $F_D > 5$, há indícios de falta de convergência (RAFTERY; LEWIS, 1992).

3 MATERIAL E MÉTODOS

Os dados foram cedidos pelo pesquisador Paulo Martins Soares Filho do Laboratório Nacional Agropecuário (LANAGRO). Trata-se de uma amostra obtida a partir de material encaminhado pelo SIF (Serviço de Inspeção Federal) ao Laboratório de Diagnóstico de Doenças Bacterianas do LANAGRO - MG, entre os anos de 2008 a 2013. Os dados são provenientes dos resultados de quatro testes diagnósticos, sendo três sorológicos: Antígeno Acidificado Tamponado (AAT), 2-Mercaptoetanol (2-ME), Fixação de Complemento (FC) e um bacteriológico (DBac), aplicados em amostras de soros sanguíneos e carcaças de animais com suspeita de brucelose bovina.

Para a realização dos testes sorológicos foram coletadas amostras de sangue que cobriram no mínimo 50% da capacidade de um tubo de 10 ml, os quais foram mantidos à temperatura ambiente por no mínimo 2 horas, ao abrigo da luz, até que ocorresse a coagulação sanguínea. Após a separação do coágulo, os soros foram transferidos para frascos limpos e secos e refrigerados até serem testados no laboratório para verificar a presença de anticorpos contra *B. abortus*.

Para realizar o teste bacteriológico, a partir das carcaças dos animais, as amostras são constituídas de: gânglios linfáticos mamários, ilíacos internos, retrofaríngeo, parótida, preescapular e baço; fetos abortados (conteúdo estomacal, baço, pulmão e fígado); útero, úbere e testículos. Todas as amostras foram embaladas separadamente e transportadas imediatamente para o laboratório e refrigerados entre 2 e 8 °C ou congeladas a -20 °C em recipientes estanques até o seu processamento. Para o isolamento e identificação, as amostras foram descongeladas em temperatura ambiente e maceradas usando um moedor de tecidos com solução salina tamponada com fosfato estéril (PBS). Após esse procedimento as amostras foram incubadas em uma atmosfera de 5% a 10% de CO₂ a 37 °C, e durante 14 dias, eram observadas diariamente para constatar se houve crescimento da bactéria.

Os testes AAT, 2-ME, FC e DBac foram realizados conforme as recomendações do Manual Técnico do Programa Nacional de Controle e Erradicação da Brucelose e Tuberculose (BRASIL, 2006).

Os resultados para os quatro testes diagnósticos: AAT = T_1 , 2-ME = T_2 , FC = T_3 e DBac = T_4 são denotados como negativo, $T_j = 0$, para os animais classificados como não doentes ou positivo, $T_j = 1$, para os animais classificados como doentes, tal que $j = 1, 2, 3, 4$. Os resultados e suas respectivas frequências observadas estão apresentados na Tabela 2.

Tabela 2: Resultados dos testes diagnósticos AAT, 2-ME, FC e DBac e suas respectivas frequências observadas.

AAT	2-ME	FC	DBac	Frequências observadas
1	1	1	1	65
1	1	1	0	100
1	1	0	1	16
1	1	0	0	15
1	0	1	1	12
1	0	1	0	12
1	0	0	1	2
1	0	0	0	4
0	1	1	1	50
0	1	1	0	85
0	1	0	1	10
0	1	0	0	17
0	0	1	1	16
0	0	1	0	23
0	0	0	1	2
0	0	0	0	6

A partir dos resultados da Tabela 2, para avaliar o desempenho dos testes diagnósticos, estimou-se o vetor de parâmetros $\theta = (\delta, S_j, E_j)$, em que δ é a prevalência da doença; S_j e E_j são as sensibilidades e especificidades de cada teste diagnóstico, respectivamente. Considerou-se inicialmente a suposição de independência condicional entre os testes diagnósticos (seção 3.1.2). Porém como na prática essa suposição pode não ser verdadeira, posteriormente, considerou-se dependência condicional entre os testes diagnósticos, (seção 3.2), de modo que o vetor θ foi acrescido de mais dois parâmetros, $\theta = (\delta, S_j, E_j, covs_{jj'}, cove_{jj'})$ com $j \neq j'$, sendo $covs_{jj'}$ a covariância entre os testes diagnósticos dado a presença da doença e $cove_{jj'}$ a covariância entre os testes dado a ausência da doença.

3.1 Modelos para testes diagnósticos considerando a suposição de independência condicional e ausência do teste Padrão Ouro

3.1.1 Um teste diagnóstico

Considere os resultados apresentados na Tabela 3 para um teste diagnóstico, elaborada conforme Joseph, Gyorkos e Coupal (1995):

Tabela 3: Possíveis resultados observados e latentes para um teste diagnóstico, considerando a ausência do teste Padrão Ouro.

Teste	Estado da doença		Total
	$D = 1$	$D = 0$	
$T = 1$	Y_1	$a - Y_1$	a
$T = 0$	Y_2	$b - Y_2$	b
Total	$Y_1 + Y_2$	$n - (Y_1 + Y_2)$	n

em que, Y_1 e Y_2 são as variáveis latentes, sendo Y_1 o número desconhecido de animais doentes entre os a classificados como positivos e Y_2 o número desconhecido de animais doentes entre os b classificados como negativos; a e b , respectivamente, são os dados observados dos animais classificados como positivos ou negativos pelo teste, aplicado em uma amostra de ($n = a + b$) animais (JOSEPH; GYORKOS; COUPAL, 1995).

De acordo com Joseph, Gyorkos e Coupal (1995) e Christensen et al. (2011), a Tabela 3 exhibe quatro resultados possíveis. Utilizando as definições fornecidas na seção 2.2 e conforme o conceito de probabilidade condicional dada em DeGroot e Schervish (2012), tem-se que:

$$\begin{aligned}
 P(T = 1, D = 1) &= P(D = 1)P(T = 1|D = 1) = \delta S_1; \\
 P(T = 0, D = 1) &= P(D = 1)P(T = 0|D = 1) = \delta(1 - S_1); \\
 P(T = 1, D = 0) &= P(D = 0)P(T = 1|D = 0) = (1 - \delta)(1 - E_1); \\
 P(T = 0, D = 0) &= P(D = 0)P(T = 0|D = 0) = (1 - \delta)E_1.
 \end{aligned}$$

Observa-se na Tabela 3 que existem duas possibilidades de ocorrências: o teste classifica o animal como positivo (sucesso) ou negativo (fracasso), sendo o animal portador ou não da doença (CHRISTENSEN et al., 2011). Assim, a partir dos resultados acima, a probabilidade do animal ser classificado como positivo é:

$$\begin{aligned}
 P(T = 1) &= P(T = 1, D = 1) + P(T = 1, D = 0) \\
 &= \delta S_1 + (1 - \delta)(1 - E_1),
 \end{aligned}$$

(BRANSCUM; GARDNER; JOHNSON, 2005). A probabilidade do animal ser classificado como negativo é dada por:

$$\begin{aligned}
 P(T = 0) &= P(T = 0, D = 1) + P(T = 0, D = 0) \\
 &= \delta(1 - S_1) + (1 - \delta)E_1.
 \end{aligned}$$

Nesse caso, o número total de animais classificados como positivos pelo teste em n

animais amostrados segue uma distribuição binomial, $Bin(n, \delta S_1 + (1 - \delta)(1 - E_1))$ (CHRISTENSEN et al., 2011). Portanto, a função de verossimilhança dos dados observados para a amostra de n animais é proporcional a distribuição binomial, isto é:

$$L(\boldsymbol{\theta}|n) \propto [\delta S_1 + (1 - \delta)(1 - E_1)]^a [\delta(1 - S_1) + (1 - \delta)E_1]^{n-a}, \quad (19)$$

conforme descreve Christensen et al. (2011).

As variáveis latentes Y_1 e Y_2 são binomialmente distribuídas, ou seja, $Y_1 \sim Bin(a, p_1)$ e $Y_2 \sim Bin(b, p_2)$ (JOSEPH; GYORKOS; COUPAL, 1995; CHRISTENSEN et al., 2011). Sabendo que cada teste fornece apenas dois resultados, pela definição de variável aleatória binomial, tem-se que, $Y_1 = \sum_{i=1}^a y_{1i}$ e $Y_2 = \sum_{i=1}^b y_{2i}$, tal que $n = a + b$. As probabilidades de sucesso, conforme Joseph, Gyorkos e Coupal (1995) são dadas por:

$$\begin{aligned} p_1 &= P(D = 1|T = 1) = \frac{P(D = 1, T = 1)}{P(T = 1)} = \frac{\delta S_1}{\delta S_1 + (1 - \delta)(1 - E_1)}; \\ p_2 &= P(D = 1|T = 0) = \frac{P(D = 1, T = 0)}{P(T = 0)} = \frac{\delta(1 - S_1)}{\delta(1 - S_1) + (1 - \delta)E_1}. \end{aligned}$$

As probabilidades de fracasso são:

$$\begin{aligned} (1 - p_1) &= P(D = 0|T = 1) = \frac{P(D = 0, T = 1)}{P(T = 1)} = \frac{(1 - \delta)(1 - E_1)}{\delta S_1 + (1 - \delta)(1 - E_1)}; \\ (1 - p_2) &= P(D = 0|T = 0) = \frac{P(D = 0, T = 0)}{P(T = 0)} = \frac{(1 - \delta)E_1}{\delta(1 - S_1) + (1 - \delta)E_1}. \end{aligned}$$

Assim, a função de verossimilhança dos dados latentes é proporcional a multiplicação das duas distribuições binomiais de Y_1 e Y_2 , como apresenta-se a seguir:

$$\begin{aligned} L(\boldsymbol{\theta}|Y_1, Y_2) &\propto \{[\delta S_1]^{Y_1} [(1 - \delta)(1 - E_1)]^{a-Y_1}\} \{[\delta S_1 + (1 - \delta)(1 - E_1)]^{-a}\} \\ &\quad \times \{[\delta(1 - S_1)]^{Y_2} [(1 - \delta)E_1]^{b-Y_2}\} \{[\delta(1 - S_1) + (1 - \delta)E_1]^{-b}\}. \end{aligned} \quad (20)$$

Multiplicando as expressões (19) e (20), como sugerido por Joseph, Gyorkos e Coupal (1995), obtém-se a função de verossimilhança aumentada, isto é:

$$L(\boldsymbol{\theta}|n, Y_1, Y_2) \propto \delta^{Y_1+Y_2} (1 - \delta)^{n-Y_1-Y_2} S_1^{Y_1} (1 - S_1)^{Y_2} E_1^{b-Y_2} (1 - E_1)^{a-Y_1}. \quad (21)$$

Segundo Jones et al. (2010), o modelo para um teste diagnóstico, $J = 1$, não apresenta condição básica para identificabilidade, pois o número de graus de liberdade é menor que o número de parâmetros, isto é, $gl = 2^J - 1 = 1 < r = 2J + 1 = 3$.

Como mostrado na seção 3.3, foi considerada a distribuição beta a priori para o vetor

de parâmetros $\boldsymbol{\theta}$, isto é, $\delta \sim \text{Beta}(\alpha_\delta, \beta_\delta)$, $S_1 \sim \text{Beta}(\alpha_{S_1}, \beta_{S_1})$ e $E_1 \sim \text{Beta}(\alpha_{E_1}, \beta_{E_1})$. Assumindo independência entre os parâmetros, a distribuição a priori conjunta é descrita como o produto das densidades marginais a priori de cada parâmetro (DENDUKURI; BÉLISLE; JOSEPH, 2010; PEREIRA et al., 2012). Então, a distribuição a priori conjunta é dada por:

$$\begin{aligned} \pi(\boldsymbol{\theta}) &\propto \pi(\delta)\pi(S_1)\pi(E_1) \\ &\propto \delta^{\alpha_\delta-1}(1-\delta)^{\beta_\delta-1}S_1^{\alpha_{S_1}-1}(1-S_1)^{\beta_{S_1}-1}E_1^{\alpha_{E_1}-1}(1-E_1)^{\beta_{E_1}-1}. \end{aligned} \quad (22)$$

Multiplicando as expressões (21) e (22), tem-se que a distribuição conjunta a posteriori é dada por:

$$\begin{aligned} \pi(\boldsymbol{\theta}|n, Y_1, Y_2) &\propto L(\boldsymbol{\theta}|n, Y_1, Y_2)\pi(\boldsymbol{\theta}) \\ &\propto \delta^{Y_1+Y_2+\alpha_\delta-1}(1-\delta)^{n-(Y_1+Y_2)+\beta_\delta-1}S_1^{Y_1+\alpha_{S_1}-1}(1-S_1)^{Y_2+\beta_{S_1}-1} \\ &\quad \times E_1^{b-Y_2+\alpha_{E_1}-1}(1-E_1)^{a-Y_1+\beta_{E_1}-1}. \end{aligned} \quad (23)$$

3.1.2 Quatro testes diagnósticos

A Tabela 4 apresenta uma configuração para os resultados observáveis e latentes para quatro testes diagnósticos:

Tabela 4: Possíveis resultados observados e latentes para quatro testes diagnósticos, considerando a ausência do teste Padrão Ouro e supondo independência condicional entre os testes.

Testes	Estado da doença		Total			
	$D = 1$	$D = 0$				
$T_1 = 1$	$T_2 = 1$	$T_3 = 1$	$T_4 = 1$	Y_1	$a - Y_1$	a
		$T_3 = 1$	$T_4 = 0$	Y_2	$b - Y_2$	b
		$T_3 = 0$	$T_4 = 1$	Y_3	$c - Y_3$	c
		$T_3 = 0$	$T_4 = 0$	Y_4	$d - Y_4$	d
	$T_2 = 0$	$T_3 = 1$	$T_4 = 1$	Y_5	$e - Y_5$	e
		$T_3 = 1$	$T_4 = 0$	Y_6	$f - Y_6$	f
		$T_3 = 0$	$T_4 = 1$	Y_7	$g - Y_7$	g
		$T_3 = 0$	$T_4 = 0$	Y_8	$h - Y_8$	h
$T_1 = 0$	$T_2 = 1$	$T_3 = 1$	$T_4 = 1$	Y_9	$k - Y_9$	k
		$T_3 = 1$	$T_4 = 0$	Y_{10}	$l - Y_{10}$	l
		$T_3 = 0$	$T_4 = 1$	Y_{11}	$m - Y_{11}$	m
		$T_3 = 0$	$T_4 = 0$	Y_{12}	$o - Y_{12}$	o
	$T_2 = 0$	$T_3 = 1$	$T_4 = 1$	Y_{13}	$q - Y_{13}$	q
		$T_3 = 1$	$T_4 = 0$	Y_{14}	$s - Y_{14}$	s
		$T_3 = 0$	$T_4 = 1$	Y_{15}	$u - Y_{15}$	u
		$T_3 = 0$	$T_4 = 0$	Y_{16}	$v - Y_{16}$	v
Total					n	

Y_1 é o número desconhecido de animais portadores da doença entre os a animais classificados como positivos em todos os testes; Y_2 o número desconhecido de animais portadores da doença entre os b animais classificados como positivo pelos testes T_1, T_2, T_3 e negativo pelo teste T_4 , e assim por diante. a, b, \dots, v são os dados observados dos animais classificados como positivos ou negativos pelos testes em uma amostra de tamanho n , tal que, $n = a + b + \dots + v$.

Para quatro testes diagnósticos, existem 32 possíveis resultados, como apresenta-se na Tabela 5,

Tabela 5: Possíveis combinações com as respectivas probabilidades para quatro testes diagnósticos, considerando-se ausência do teste Padrão Ouro e supondo-se independência condicional entre os testes.

Probabilidades	Resultados
$P(T_1 = 1, T_2 = 1, T_3 = 1, T_4 = 1, D = 1)$	$\delta S_1 S_2 S_3 S_4$
$P(T_1 = 1, T_2 = 1, T_3 = 1, T_4 = 0, D = 1)$	$\delta S_1 S_2 S_3 (1 - S_4)$
$P(T_1 = 1, T_2 = 1, T_3 = 0, T_4 = 1, D = 1)$	$\delta S_1 S_2 (1 - S_3) S_4$
$P(T_1 = 1, T_2 = 1, T_3 = 0, T_4 = 0, D = 1)$	$\delta S_1 S_2 (1 - S_3) (1 - S_4)$
$P(T_1 = 1, T_2 = 0, T_3 = 1, T_4 = 1, D = 1)$	$\delta S_1 (1 - S_2) S_3 S_4$
$P(T_1 = 1, T_2 = 0, T_3 = 1, T_4 = 0, D = 1)$	$\delta S_1 (1 - S_2) S_3 (1 - S_4)$
$P(T_1 = 1, T_2 = 0, T_3 = 0, T_4 = 1, D = 1)$	$\delta S_1 (1 - S_2) (1 - S_3) S_4$
$P(T_1 = 1, T_2 = 0, T_3 = 0, T_4 = 0, D = 1)$	$\delta S_1 (1 - S_2) (1 - S_3) (1 - S_4)$
$P(T_1 = 0, T_2 = 1, T_3 = 1, T_4 = 1, D = 1)$	$\delta (1 - S_1) S_2 S_3 S_4$
$P(T_1 = 0, T_2 = 1, T_3 = 1, T_4 = 0, D = 1)$	$\delta (1 - S_1) S_2 S_3 (1 - S_4)$
$P(T_1 = 0, T_2 = 1, T_3 = 0, T_4 = 1, D = 1)$	$\delta (1 - S_1) S_2 (1 - S_3) S_4$
$P(T_1 = 0, T_2 = 1, T_3 = 0, T_4 = 0, D = 1)$	$\delta (1 - S_1) S_2 (1 - S_3) (1 - S_4)$
$P(T_1 = 0, T_2 = 0, T_3 = 1, T_4 = 1, D = 1)$	$\delta (1 - S_1) (1 - S_2) S_3 S_4$
$P(T_1 = 0, T_2 = 0, T_3 = 1, T_4 = 0, D = 1)$	$\delta (1 - S_1) (1 - S_2) S_3 (1 - S_4)$
$P(T_1 = 0, T_2 = 0, T_3 = 0, T_4 = 1, D = 1)$	$\delta (1 - S_1) (1 - S_2) (1 - S_3) S_4$
$P(T_1 = 0, T_2 = 0, T_3 = 0, T_4 = 0, D = 1)$	$\delta (1 - S_1) (1 - S_2) (1 - S_3) (1 - S_4)$
$P(T_1 = 1, T_2 = 1, T_3 = 1, T_4 = 1, D = 0)$	$(1 - \delta) (1 - E_1) (1 - E_2) (1 - E_3) (1 - E_4)$
$P(T_1 = 1, T_2 = 1, T_3 = 1, T_4 = 0, D = 0)$	$(1 - \delta) (1 - E_1) (1 - E_2) (1 - E_3) E_4$
$P(T_1 = 1, T_2 = 1, T_3 = 0, T_4 = 1, D = 0)$	$(1 - \delta) (1 - E_1) (1 - E_2) E_3 (1 - E_4)$
$P(T_1 = 1, T_2 = 1, T_3 = 0, T_4 = 0, D = 0)$	$(1 - \delta) (1 - E_1) (1 - E_2) E_3 E_4$
$P(T_1 = 1, T_2 = 0, T_3 = 1, T_4 = 1, D = 0)$	$(1 - \delta) (1 - E_1) E_2 (1 - E_3) (1 - E_4)$
$P(T_1 = 1, T_2 = 0, T_3 = 1, T_4 = 0, D = 0)$	$(1 - \delta) (1 - E_1) E_2 (1 - E_3) E_4$
$P(T_1 = 1, T_2 = 0, T_3 = 0, T_4 = 1, D = 0)$	$(1 - \delta) (1 - E_1) E_2 E_3 (1 - E_4)$
$P(T_1 = 1, T_2 = 0, T_3 = 0, T_4 = 0, D = 0)$	$(1 - \delta) (1 - E_1) E_2 E_3 E_4$
$P(T_1 = 0, T_2 = 1, T_3 = 1, T_4 = 1, D = 0)$	$(1 - \delta) E_1 (1 - E_2) (1 - E_3) (1 - E_4)$
$P(T_1 = 0, T_2 = 1, T_3 = 1, T_4 = 0, D = 0)$	$(1 - \delta) E_1 (1 - E_2) (1 - E_3) E_4$
$P(T_1 = 0, T_2 = 1, T_3 = 0, T_4 = 1, D = 0)$	$(1 - \delta) E_1 (1 - E_2) E_3 (1 - E_4)$
$P(T_1 = 0, T_2 = 1, T_3 = 0, T_4 = 0, D = 0)$	$(1 - \delta) E_1 (1 - E_2) E_3 E_4$
$P(T_1 = 0, T_2 = 0, T_3 = 1, T_4 = 1, D = 0)$	$(1 - \delta) E_1 E_2 (1 - E_3) (1 - E_4)$
$P(T_1 = 0, T_2 = 0, T_3 = 1, T_4 = 0, D = 0)$	$(1 - \delta) E_1 E_2 (1 - E_3) E_4$
$P(T_1 = 0, T_2 = 0, T_3 = 0, T_4 = 1, D = 0)$	$(1 - \delta) E_1 E_2 E_3 (1 - E_4)$
$P(T_1 = 0, T_2 = 0, T_3 = 0, T_4 = 0, D = 0)$	$(1 - \delta) E_1 E_2 E_3 E_4$

sendo $P(T_1, T_2, T_3, T_4, D) = P(D)P(T_1|D)P(T_2|D)P(T_3|D)P(T_4|D)$.

Por meio da Tabela 5, verifica-se que existem dezesseis alternativas para os resultados dos quatro testes diagnósticos, sendo o animal portador ou não da doença. Assim os dados

observados dos testes diagnósticos (a, b, \dots, v) indicam o número de vezes que ocorre cada alternativa de resultados nos n animais amostrados. Dessa forma, os dados observados dos testes seguem uma distribuição multinomial, $Mult(n, (p_a, p_b, p_c, \dots, p_v))$, e as probabilidades de ocorrências, conforme os resultados acima, são resumidamente dadas por:

$$\begin{aligned}
p_a &= P(T_1 = 1, T_2 = 1, T_3 = 1, T_4 = 1, D = 1) + P(T_1 = 1, T_2 = 1, T_3 = 1, T_4 = 1, D = 0) \\
&= \delta S_1 S_2 S_3 S_4 + (1 - \delta)(1 - E_1)(1 - E_2)(1 - E_3)(1 - E_4); \\
p_b &= P(T_1 = 1, T_2 = 1, T_3 = 1, T_4 = 0, D = 1) + P(T_1 = 1, T_2 = 1, T_3 = 1, T_4 = 0, D = 0) \\
&= \delta S_1 S_2 S_3(1 - S_4) + (1 - \delta)(1 - E_1)(1 - E_2)(1 - E_3)E_4; \\
p_c &= P(T_1 = 1, T_2 = 1, T_3 = 0, T_4 = 1, D = 1) + P(T_1 = 1, T_2 = 1, T_3 = 0, T_4 = 1, D = 0) \\
&= \delta S_1 S_2(1 - S_3)S_4 + (1 - \delta)(1 - E_1)(1 - E_2)E_3(1 - E_4); \\
&\vdots \\
p_v &= P(T_1 = 0, T_2 = 0, T_3 = 0, T_4 = 0, D = 1) + P(T_1 = 0, T_2 = 0, T_3 = 0, T_4 = 0, D = 0) \\
&= \delta(1 - S_1)(1 - S_2)(1 - S_3)(1 - S_4) + (1 - \delta)E_1 E_2 E_3 E_4.
\end{aligned}$$

Assim, a função de verossimilhança dos dados observados é proporcional a distribuição multinomial dada a seguir:

$$\begin{aligned}
L(\boldsymbol{\theta}|n) &\propto [\delta S_1 S_2 S_3 S_4 + (1 - \delta)(1 - E_1)(1 - E_2)(1 - E_3)(1 - E_4)]^a \\
&\quad \times \dots \times [\delta(1 - S_1)(1 - S_2)(1 - S_3)(1 - S_4) + (1 - \delta)E_1 E_2 E_3 E_4]^v. \quad (24)
\end{aligned}$$

Para quatro testes diagnósticos, as variáveis latentes: $Y_1 = \sum_{i=1}^a y_{1i}$, $Y_2 = \sum_{i=1}^b y_{2i}$, $Y_3 = \sum_{i=1}^c y_{3i}$, \dots , $Y_{16} = \sum_{i=1}^v y_{16i}$, apresentam as seguintes distribuições binomiais: $Y_1 \sim Bin(a, p_1)$, $Y_2 \sim Bin(b, p_2)$, $Y_3 \sim Bin(c, p_3)$, \dots , $Y_{16} \sim Bin(v, p_{16})$, sendo as probabilidades de sucesso dadas por:

$$\begin{aligned}
p_1 &= \frac{\delta S_1 S_2 S_3 S_4}{\delta S_1 S_2 S_3 S_4 + (1 - \delta)(1 - E_1)(1 - E_2)(1 - E_3)(1 - E_4)}; \\
p_2 &= \frac{\delta S_1 S_2 S_3(1 - S_4)}{\delta S_1 S_2 S_3(1 - S_4) + (1 - \delta)(1 - E_1)(1 - E_2)(1 - E_3)E_4}; \\
p_3 &= \frac{\delta S_1 S_2(1 - S_3)S_4}{\delta S_1 S_2(1 - S_3)S_4 + (1 - \delta)(1 - E_1)(1 - E_2)E_3(1 - E_4)}; \\
&\vdots \\
p_{16} &= \frac{\delta(1 - S_1)(1 - S_2)(1 - S_3)(1 - S_4)}{\delta(1 - S_1)(1 - S_2)(1 - S_3)(1 - S_4) + (1 - \delta)E_1 E_2 E_3 E_4}.
\end{aligned}$$

Logo, a função de verossimilhança dos dados latentes é proporcional a multiplicação das

dezesseis distribuições binomiais, sendo as probabilidades de fracasso obtidas de forma similar as apresentadas para o caso de um teste diagnóstico, tem-se que:

$$L(\boldsymbol{\theta}|Y_1, \dots, Y_{16}) \propto \left[\frac{\delta S_1 S_2 S_3 S_4}{\delta S_1 S_2 S_3 S_4 + (1 - \delta)(1 - E_1)(1 - E_2)(1 - E_3)(1 - E_4)} \right]^{Y_1} \\ \times \dots \times \left[\frac{(1 - \delta)E_1 E_2 E_3 E_4}{\delta(1 - S_1)(1 - S_2)(1 - S_3)(1 - S_4) + (1 - \delta)E_1 E_2 E_3 E_4} \right]^{v - Y_{16}}. \quad (25)$$

Combinando as expressões (24) e (25), a função de verossimilhança aumentada é:

$$L(\boldsymbol{\theta}|n, Y_1, \dots, Y_{16}) \propto \delta^{Y_1 + \dots + Y_{16}} (1 - \delta)^{n - (Y_1 + \dots + Y_4)} S_1^{Y_1 + \dots + Y_8} (1 - S_1)^{Y_9 + \dots + Y_{16}} \times \dots \\ \times (1 - E_4)^{(a+c+e+g+k+m+q+u) - (Y_1 + Y_3 + Y_5 + Y_7 + Y_9 + Y_{11} + Y_{13} + Y_{15})}. \quad (26)$$

O modelo para quatro testes diagnósticos, $J = 4$, apresenta condição básica de identificabilidade, uma vez que, $gl = 2^J - 1 = 15 > r = 2J + 1 = 9$.

Assumindo distribuição beta a priori para o vetor $\boldsymbol{\theta} = (\delta, S_1, S_2, S_3, S_4, E_1, E_2, E_3, E_4)$ e considerando independência entre os parâmetros (DENDUKURI; BÉLISLE; JOSEPH, 2010; PEREIRA et al., 2012), tem-se que a distribuição conjunta a priori para os parâmetros de quatro testes diagnósticos é apresentada por:

$$\pi(\boldsymbol{\theta}) \propto \pi(\delta)\pi(S_1)\pi(S_2)\pi(S_3)\pi(S_4)\pi(E_1)\pi(E_2)\pi(E_3)\pi(E_4) \\ \propto \delta^{\alpha_\delta - 1} (1 - \delta)^{\beta_\delta - 1} \prod_{j=1}^4 S_j^{\alpha_{S_j} - 1} (1 - S_j)^{\beta_{S_j} - 1} E_j^{\alpha_{E_j} - 1} (1 - E_j)^{\beta_{E_j} - 1}. \quad (27)$$

E por fim, a distribuição conjunta a posteriori é obtida combinando as expressões (26) e (27), assim,

$$\pi(\boldsymbol{\theta}|n, Y_1, \dots, Y_{16}) \propto L(\boldsymbol{\theta}|n, Y_1, \dots, Y_{16})\pi(\boldsymbol{\theta}) \\ \propto \delta^{Y_1 + \dots + Y_{16}} (1 - \delta)^{n - (Y_1 + \dots + Y_4)} S_1^{Y_1 + \dots + Y_8} (1 - S_1)^{Y_9 + \dots + Y_{16}} \times \dots \\ \times (1 - E_4)^{(a+c+e+g+k+m+q+u) - (Y_1 + Y_3 + Y_5 + Y_7 + Y_9 + Y_{11} + Y_{13} + Y_{15})} \times \\ \times \delta^{\alpha_\delta - 1} (1 - \delta)^{\beta_\delta - 1} \prod_{j=1}^4 S_j^{\alpha_{S_j} - 1} (1 - S_j)^{\beta_{S_j} - 1} E_j^{\alpha_{E_j} - 1} (1 - E_j)^{\beta_{E_j} - 1}. \quad (28)$$

3.2 Modelos para testes diagnósticos considerando dependência condicional e ausência do teste Padrão Ouro

3.2.1 Quatro testes diagnósticos

Como descrito na seção 2.3.1, os testes AAT, 2-ME e FC podem ser condicionalmente dependentes, sendo o teste DBac condicionalmente independente dos demais testes. Se dois testes são condicionalmente dependentes, $P(T_1 = 1, T_2 = 1|D = 1) \neq S_1 S_2$ e/ou $P(T_1 = 0, T_2 = 0|D = 0) \neq E_1 E_2$; se são condicionalmente independentes, $P(T_1 = 1, T_2 = 1|D = 1) = S_1 S_2$ e $P(T_1 = 0, T_2 = 0|D = 0) = E_1 E_2$ (GARDNER et al., 2000). Para avaliar as possíveis dependências entre os testes em estudo, a abordagem sugerida por Vacek (1985) foi adotada. Essa estratégia permite que covariâncias sejam incorporadas entre os testes diagnósticos dado o estado da doença.

Geralmente, a dependência condicional entre testes diagnósticos é avaliada dois a dois, a medida que o número de testes aumenta, mais complicada torna-se a notação, principalmente se dependências de alta ordem são incorporadas (MENTEN; BOELAERT; LESAFFRE, 2008), por exemplo, se a dependência condicional é analisada três a três, para mais detalhes ver (ADEL; BERKVENS, 2002). Apesar de serem possíveis, dependências de ordem superior frequentemente não ocorrem na prática, além disso, a interpretação é comprometida (TORRANCE-RYNARD; WALTER, 1997). Portanto, para este trabalho, a análise foi conduzida incorporando pares de covariâncias entre os testes diagnósticos AAT, 2-ME e FC. A dependência condicional entre os testes diagnósticos foi avaliada considerando os seguintes cenários: dependência entre os testes AAT e 2-ME, dependência entre os testes AAT e FC, dependência entre os testes 2-ME e FC, por fim, considerando todos os pares de dependência (AAT e 2-ME, AAT e FC, 2-ME e FC).

Considerando os testes diagnósticos AAT e 2-ME, a covariância condicional entre dois testes é dada por:

$$cov(T_1, T_2|D) = E(T_1.T_2|D) - E(T_1|D)E(T_2|D). \quad (29)$$

De acordo com Dendukuri e Joseph (2001), dois testes diagnósticos apresentam os resultados: $(T_1 = 1, T_2 = 1)$, $(T_1 = 1, T_2 = 0)$, $(T_1 = 0, T_2 = 1)$ e $(T_1 = 0, T_2 = 0)$, sendo o animal portador ou não da doença. Conforme Gardner et al. (2000), os valores esperados para os resultados de dois testes diagnósticos, por exemplo, o resultado $(T_1 = 1, T_2 = 1)$, dado que o animal é portador da doença, são dados da seguinte maneira:

$$\begin{aligned}
E(T_1.T_2|D = 1) &= \sum_{t_1, t_2} t_1 t_2 P(T_1 = t_1, T_2 = t_2 | D = 1) \\
&= 1.1 P(T_1 = 1, T_2 = 1 | D = 1) + \dots + 0.0 P(T_1 = 0, T_2 = 0 | D = 1) \\
&= P(T_1 = 1, T_2 = 1 | D = 1).
\end{aligned}$$

Aplicando o mesmo raciocínio,

$$\begin{aligned}
E(T_1|D = 1) &= \sum_{t_1} t_1 P(T_1 = t_1 | D = 1) \\
&= 1. P(T_1 = 1 | D = 1) + 0. P(T_1 = 0 | D = 1) \\
&= P(T_1 = 1 | D = 1) \\
&= S_1.
\end{aligned}$$

e

$$\begin{aligned}
E(T_2|D = 1) &= \sum_{t_2} t_2 P(T_2 = t_2 | D = 1) \\
&= 1. P(T_2 = 1 | D = 1) + 0. P(T_2 = 0 | D = 1) \\
&= P(T_2 = 1 | D = 1) \\
&= S_2.
\end{aligned}$$

Supondo o mesmo procedimento para as demais resultados, a expressão (29), de forma geral, pode ser reescrita como:

$$cov(T_1, T_2 | D) = P(T_1, T_2 | D) - P(T_1 | D)P(T_2 | D). \quad (30)$$

Com base na expressão acima e considerando os resultados entre os quatro testes diagnósticos apresentadas na Tabela 5, para incorporar a covariância condicional, suponha o seguinte resultado: $(T_1 = 1, T_2 = 1, T_3 = 1, T_4 = 1, D = 1)$. De acordo com Dendukuri e Joseph (2001), a covariância condicional entre dois testes, é dada da seguinte maneira:

$$\begin{aligned}
cous_{12} &= cov(T_1 = 1, T_2 = 1 | D = 1) \\
&= P(T_1 = 1, T_2 = 1 | D = 1) - P(T_1 = 1 | D = 1)P(T_2 = 1 | D = 1) \\
&= P(T_1 = 1, T_2 = 1 | D = 1) - S_1 S_2,
\end{aligned} \quad (31)$$

sendo $cous_{12}$ a covariância entre os testes diagnósticos dado que o animal é portador da doença (GARDNER et al., 2000; BRANSCUM; GARDNER; JOHNSON, 2005). Sabendo que,

$$\begin{aligned}
P(T_1 = 1, T_2 = 1 | D = 1) &= S_1 S_2 + cov_{S_{12}} \\
\frac{P(T_1 = 1, T_2 = 1, D = 1)}{P(D = 1)} &= S_1 S_2 + cov_{S_{12}} \\
P(T_1 = 1, T_2 = 1, D = 1) &= \delta[S_1 S_2 + cov_{S_{12}}],
\end{aligned}$$

então, a proporção esperada para o resultado, com a covariância condicional incorporada, é dado da seguinte forma: $P(T_1 = 1, T_2 = 1, T_3 = 1, T_4 = 1, D = 1) = \delta[S_1 S_2 + cov_{S_{12}}] S_3 S_4$.

De maneira semelhante, para o resultado, $(T_1 = 0, T_2 = 0, T_3 = 0, T_4 = 0, D = 0)$, ainda conforme Dendukuri e Joseph (2001), tem-se que:

$$\begin{aligned}
cove_{12} &= cov(T_1 = 0, T_2 = 0 | D = 0) \\
&= P(T_1 = 0, T_2 = 0 | D = 0) - P(T_1 = 0 | D = 0)P(T_2 = 0 | D = 0) \\
&= P(T_1 = 0, T_2 = 0 | D = 0) - E_1 E_2,
\end{aligned} \tag{32}$$

sendo $cove_{12}$ a covariância entre os testes diagnósticos dado que o animal não é portador da doença (GARDNER et al., 2000; BRANSCUM; GARDNER; JOHNSON, 2005). Sabendo que,

$$\begin{aligned}
P(T_1 = 0, T_2 = 0 | D = 0) &= E_1 E_2 + cove_{12} \\
\frac{P(T_1 = 0, T_2 = 0, D = 0)}{P(D = 0)} &= E_1 E_2 + cove_{12} \\
P(T_1 = 0, T_2 = 0, D = 0) &= (1 - \delta)[E_1 E_2 + cove_{12}],
\end{aligned}$$

então, a proporção esperada para o resultado, com a covariância condicional adicionada, é dada por: $P(T_1 = 0, T_2 = 0, T_3 = 0, T_4 = 0, D = 0) = (1 - \delta)[E_1 E_2 + cove_{12}] E_3 E_4$. Para as demais combinações, as covariâncias condicionais são incorporadas de forma semelhante.

Basendo-se em Dendukuri e Joseph (2001), a função de verossimilhança aumentada, para o modelo com quatro testes diagnósticos considerando os testes 1 e 2 condicionalmente dependentes e adotando o mesmo procedimento apresentado na seção 3.1.2, é dada por:

$$\begin{aligned}
L(\boldsymbol{\theta} | n, Y_1, \dots, Y_{16}) &\propto \{\delta[S_1 S_2 + cov_{S_{12}}] S_3 S_4\}^{Y_1} \\
&\times \dots \times \{(1 - \delta)[E_1 E_2 + cove_{12}] E_3 E_4\}^{v - Y_{16}}.
\end{aligned} \tag{33}$$

Segundo Jones et al. (2010), o modelo para quatro testes diagnósticos, considerando o relaxamento da suposição de independência condicional entre dois testes, apresenta condição básica de identificabilidade, pois o número de graus de liberdade ($gl = 2^J - 1 = 16 - 1 = 15$) é maior que o número de parâmetros a serem estimados ($r = 11$).

Assumindo independência entre os parâmetros do vetor $\boldsymbol{\theta} = (\delta, S_1, S_2, S_3, S_4, E_1, E_2, E_3, E_4, cov_{S_{12}}, cov_{E_{12}})$ (DENDUKURI; BÉLISLE; JOSEPH, 2010; PEREIRA et al., 2012) e considerando as distribuições a priori definidas na seção 3.3, a distribuição conjunta a priori é dada por:

$$\begin{aligned}
\pi(\boldsymbol{\theta}) &= \pi(\delta)\pi(S_1)\pi(S_2)\pi(S_3)\pi(S_4)\pi(E_1)\pi(E_2)\pi(E_3)\pi(E_4)\pi(cov_{S_{12}})\pi(cov_{E_{12}}) \\
&\propto \delta^{\alpha_\delta-1}(1-\delta)^{\beta_\delta-1} \prod_{j=1}^4 S_j^{\alpha_{S_j}-1} (1-S_j)^{\beta_{S_j}-1} E_j^{\alpha_{E_j}-1} (1-E_j)^{\beta_{E_j}-1} \\
&\times (cov_{S_{12}} - L_1)^{\alpha_{cov_{S_{12}}}-1} (L_2 - cov_{S_{12}})^{\beta_{cov_{S_{12}}}-1} \\
&\times (cov_{E_{12}} - L_1)^{\alpha_{cov_{E_{12}}}-1} (L_2 - cov_{E_{12}})^{\beta_{cov_{E_{12}}}-1}.
\end{aligned} \tag{34}$$

Combinando as expressões (33) e (34) a distribuição conjunta a posteriori é:

$$\begin{aligned}
\pi(\boldsymbol{\theta}|n, Y_1, \dots, Y_{16}) &= L(\boldsymbol{\theta}|n, Y_1, \dots, Y_{16})\pi(\boldsymbol{\theta}) \\
&\propto \{\delta[S_1S_2 + cov_{S_{12}}]S_3S_4\}^{Y_1} \times \dots \times \delta^{\alpha_\delta-1}(1-\delta)^{\beta_\delta-1} \\
&\times \prod_{j=1}^4 S_j^{\alpha_{S_j}-1} (1-S_j)^{\beta_{S_j}-1} E_j^{\alpha_{E_j}-1} (1-E_j)^{\beta_{E_j}-1} \\
&\times (cov_{S_{12}} - L_1)^{\alpha_{cov_{S_{12}}}-1} (L_2 - cov_{S_{12}})^{\beta_{cov_{S_{12}}}-1} \\
&\times (cov_{E_{12}} - L_1)^{\alpha_{cov_{E_{12}}}-1} (L_2 - cov_{E_{12}})^{\beta_{cov_{E_{12}}}-1}.
\end{aligned} \tag{35}$$

A partir desses resultados, o procedimento de obtenção da distribuição conjunta a posteriori dos parâmetros para os demais cenários de dependências condicionais é realizado de forma similar.

Conforme os parâmetros de covariâncias definidos nas expressões (31) e (32), Georgiadis et al. (2003) introduziram o conceito de correlação entre os resultados dos testes diagnósticos, isto é:

$$\begin{aligned}
\rho_D &= \frac{cov_{S_{12}}}{\sqrt{S_1(1-S_1)S_2(1-S_2)}}, \\
\rho_{ND} &= \frac{cov_{E_{12}}}{\sqrt{E_1(1-E_1)E_2(1-E_2)}},
\end{aligned}$$

sendo ρ_D , a correlação entre os resultados dos testes para os animais portadores da doença e ρ_{ND} a correlação entre os resultados dos testes para os animais não portadores da doença.

3.3 Informação a priori

Como descrito na seção 2.4, adotou-se a distribuição beta a priori para o vetor de parâmetros $\theta = (\delta, S_j, E_j)$. Para determinar os hiperparâmetros das distribuições a priori, o conhecimento do especialista sobre o vetor θ permite estabelecer um intervalo $[l_1, l_2]$ (JOSEPH; GYORKOS; COUPAL, 1995; DENDUKURI; JOSEPH, 2001). Então, a ideia é obter os hiperparâmetros: α , igualando-se as médias das distribuições beta e uniforme (SOUZA, 2013), isto é:

$$\frac{l_1 + l_2}{2} = \frac{\alpha}{\alpha + \beta}, \quad (36)$$

e β , igualando-se o desvio padrão da distribuição beta com 1/4 da amplitude do intervalo $[l_1, l_2]$ (BROWNE, 2001; RHIEL; MARKOWSKI, 2017):

$$\frac{l_2 - l_1}{4} = \sqrt{\frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}}. \quad (37)$$

A partir das igualdades (36) e (37), foram obtidas as expressões a seguir:

$$\alpha = \frac{\beta(l_1 + l_2)}{2 - l_1 - l_2} \quad \text{e} \quad \beta = \frac{2(l_1 + l_2)(2 - l_1 - l_2)^2}{(l_2 - l_1)^2} + \frac{(l_1 + l_2)}{2} - 1. \quad (38)$$

Utilizando as expressões em (38), as estimativas dos hiperparâmetros α e β para as distribuições a priori dos parâmetros S_j e E_j , com $j = 1, 2, 3, 4$, foram adquiridas, conforme os intervalos fornecidos pelo especialista do LANAGRO - MG, como apresenta-se na Tabela 6.

Tabela 6: Estimativas dos hiperparâmetros das distribuições beta para os parâmetros de sensibilidade (S), especificidade (E) para os testes AAT, 2-ME, FC e DBac obtidas conforme os intervalos $[l_1, l_2]\%$.

Parâmetros	AAT			2-ME		
	$[l_1; l_2]$	α	β	$[l_1; l_2]$	α	β
S	[95, 90; 99, 30]	315, 45	7, 75	[83, 00; 94, 00]	118, 22	15, 36
E	[99, 80; 99, 90]	4389, 33	4, 83	[89, 00; 94, 00]	454, 53	42, 22
	FC			DBac		
	$[l_1; l_2]$	α	β	$[l_1; l_2]$	α	β
S	[92, 60; 97, 50]	297, 06	15, 47	[40, 00; 60, 00]	49, 50	49, 50
E	[99, 90; 1, 00]	791, 03	3, 97	[99, 00; 1, 00]	791, 03	3, 97

Para a prevalência (δ), as estimativas dos hiperparâmetros α e β para a distribuição a priori foram $\alpha = 119, 30$ e $\beta = 51, 13$, isto é, $\delta \sim beta(\alpha = 119, 30; \beta = 51, 13)$, conforme o intervalo $[l_1, l_2]\% = [63, 00; 77, 00]$.

Como já relatado nesse texto, na prática, geralmente não existem informações a priori

para os dois parâmetros de covariâncias, nesses casos, Dendukuri e Joseph (2001), Branscum, Gardner e Johnson (2005) sugerem utilizar a distribuição uniforme a priori. Adotando a estratégia sugerida pelos autores, as covariâncias $cov_{s_{12}}$ e $cov_{e_{12}}$ variam nos intervalos:

$$(S_1 - 1)(1 - S_2) \leq cov_{s_{12}} \leq \min(S_1, S_2) - S_1 S_2 \quad (39)$$

$$(E_1 - 1)(1 - E_2) \leq cov_{e_{12}} \leq \min(E_1, E_2) - E_1 E_2. \quad (40)$$

Com isso, as distribuições uniformes a priori para as covariâncias entre os testes dada a presença ou ausência da doença são definidas como:

$$cov_{s_{12}} \sim U[(S_1 - 1)(1 - S_2), \min(S_1, S_2) - S_1 S_2],$$

$$cov_{e_{12}} \sim U[(E_1 - 1)(1 - E_2), \min(E_1, E_2) - E_1 E_2].$$

3.4 Critério de informação da deviance

Para avaliar a qualidade do ajuste dos modelos em estudo será empregado o critério de informação da deviance (DIC), proposto por Spiegelhalter et al. (2002). O DIC é uma generalização do critério de informação de Akaike (AIC), composto pela medida da qualidade do ajuste (\bar{D}) e a medida de complexidade do modelo (p_D), também denominada número efetivo de parâmetros (SPIEGELHALTER et al., 2002; BERG; MEYER; YU, 2004; ROBERT, 2007; CARLIN; LOUIS, 2008). A deviance $D(\theta) = -2\ln f(\mathbf{x}|\theta)$, sendo θ o vetor de parâmetros e $f(\mathbf{x}|\theta)$ a função de verossimilhança. A medida de complexidade do modelo é dada por:

$$p_D = \bar{D} - D(\bar{\theta}),$$

tal que, $\bar{\theta}$ é o estimador bayesiano do parâmetro θ , isto é, a média a posteriori, \bar{D} a média a posteriori da deviance e $D(\bar{\theta})$ a deviance avaliada na média a posteriori dos parâmetros, então, o DIC é definido como:

$$DIC = \bar{D} + p_D. \quad (41)$$

Como regra geral, se a diferença em DIC entre dois modelos for maior que 10, aquele com menor DIC, definitivamente, é o modelo com melhor ajuste (LUNN et al., 2012); diferenças entre 5 e 10 indicam resultados substanciais à favor do modelo com o menor DIC (STONE; ZHU, 2015) e se a diferença for menor que 5, os dois modelos são considerados semelhantes (LEE, 2007; STONE; ZHU, 2015). Ocorrendo o último caso, a escolha do modelo é realizada com base na parcimônia, ou seja, quanto menor o número efetivo de parâmetros, p_D , melhor o

ajuste do modelo aos dados (RAHMAN et al., 2013).

3.5 Procedimentos computacionais

Os modelos de classe latente na análise dos testes individuais e combinados foram implementados no programa OpenBugs 3.2.3 e executados a partir do programa R 3.2.5 (R Core Team, 2016), por meio do pacote *R2OpenBUGS* (STURTZ; LIGGES; GELMAN, 2005). O desenvolvimento dos modelos para cada teste individualmente está conforme Branscum, Gardner e Johnson (2005) e os modelos para os testes combinados, de acordo com Menten, Boelaert e Lesaffre (2008). As distribuições a posteriori dos parâmetros de interesse foram simuladas empregando os métodos de Monte Carlo via Cadeias de Markov (MCMC) (GELMAN et al., 2004) e resumidas utilizando as médias e intervalos com 95% de credibilidade (BRANSCUM; GARDNER; JOHNSON, 2005; MENTEN; BOELAERT; LESAFFRE, 2008). Os desvios padrão também foram obtidos.

Foram simuladas duas cadeias, considerando 80.000 iterações com período de aquecimento de 40.000 e espaçamento de 120 iterações. Para cada cadeia empregou-se diferentes valores iniciais. A convergência das cadeias foram avaliadas por meio dos critérios de Raftery e Lewis (1992) e Gelman e Rubin (1992), implementados no pacote *coda*. Os gráficos dos traços das cadeias foram desenvolvidos por meio da função *traceplot*. Os gráficos das densidades a posteriori e das autocorrelações foram obtidos utilizando a função *densplot* e *acf*, respectivamente.

Para avaliar o ajuste dos modelos dos testes combinados, utilizou-se o critério de informação da deviance (DIC) e o número efetivo de parâmetros (p_D).

4 RESULTADOS E DISCUSSÃO

Nesta seção, estão apresentados os resultados dos quatro testes diagnósticos avaliados individualmente e combinados. Considerando os testes analisados individualmente. Nas Figuras 1 a 4 estão apresentados os gráficos dos traços, das autocorrelações e densidades dos parâmetros: sensibilidades e especificidades dos testes diagnósticos AAT, 2-ME, FC e DBac, respectivamente, e prevalência da doença.

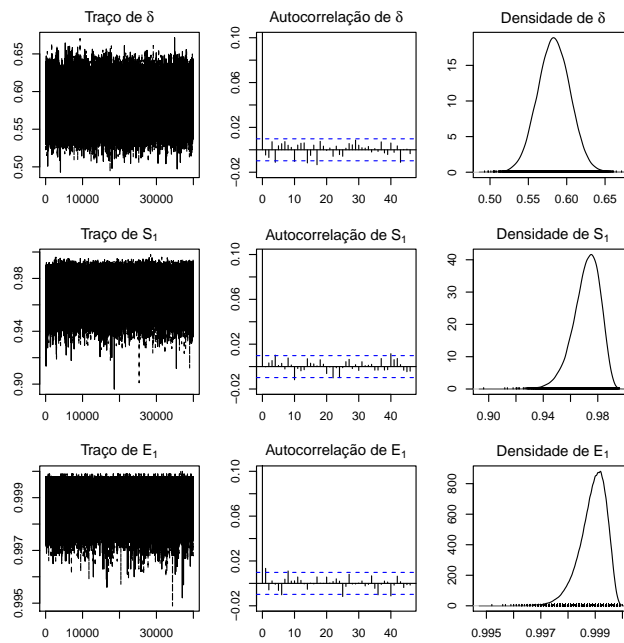


Figura 1: Gráficos dos traços, correlações e densidades dos parâmetros δ , S e E para o teste diagnóstico AAT.

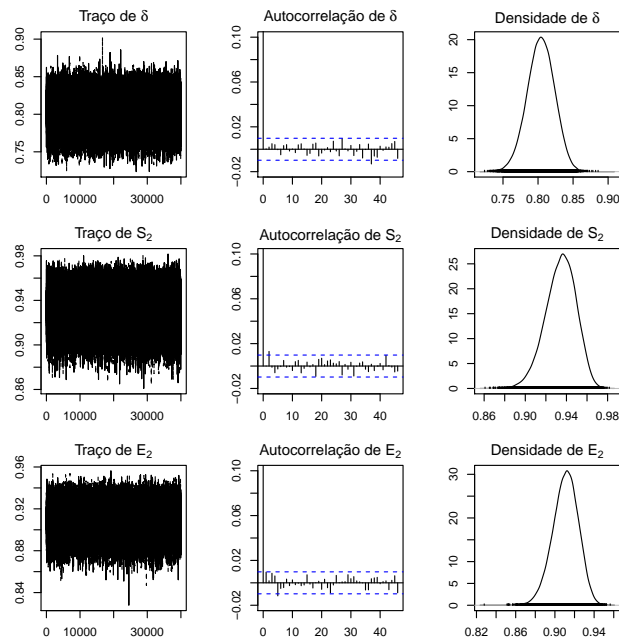


Figura 2: Gráficos dos traços, correlações e densidades dos parâmetros δ , S e E para o teste diagnóstico 2-ME.

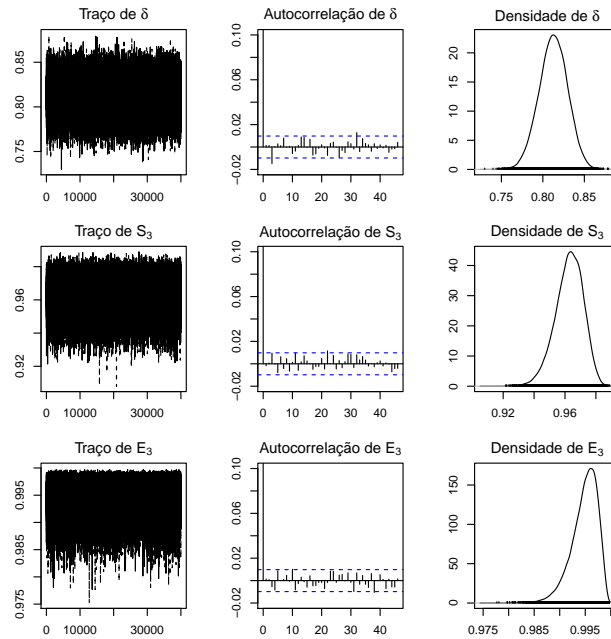


Figura 3: Gráficos dos traços, correlações e densidades dos parâmetros δ , S e E para o teste diagnóstico FC.

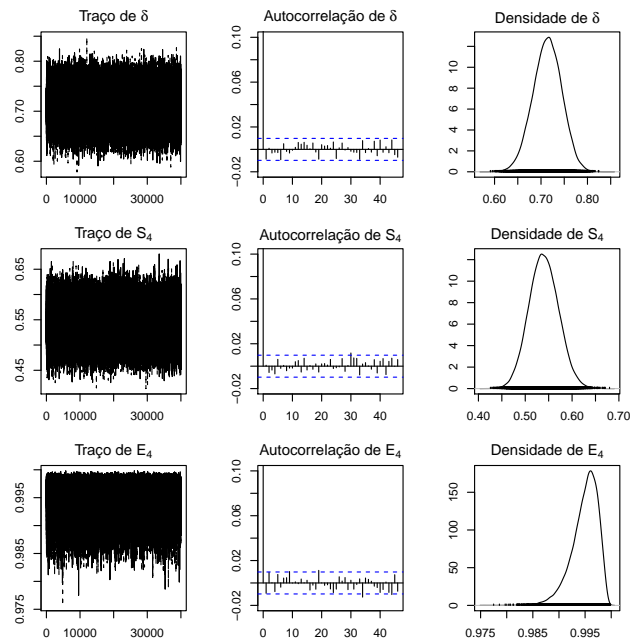


Figura 4: Gráficos dos traços, correlações e densidades dos parâmetros δ , S e E para o teste diagnóstico DBac.

Conforme os gráficos dos traços verifica-se um comportamento padrão em torno dos valores gerados dos parâmetros. Isso indica que o traçado das cadeias amostradas percorrem a mesma região do espaço dos parâmetros e convergem para a distribuição a posteriori de interesse de cada parâmetro em estudo (HAMRA; MACLEHOSE; RICHARDSON, 2013). Nos gráficos de autocorrelações, a partir do primeiro lag, os valores da autocorrelações diminuíram rapidamente conforme aumentavam os lags, de forma que os lags se concentraram dentro dos

limites de significância. Esse comportamento observado nos gráficos revela que a partir do primeiro lag, a autocorrelação é aproximadamente zero (COWLES, 2013), assim há evidências que as amostras geradas não são dependentes. Observa-se também que os gráficos das densidades apresentam simetria em torno das estimativas dos parâmetros, exceto para os parâmetros da sensibilidade (S_1) do teste AAT e as especificidades (E_1 , E_3 e E_4) dos testes diagnósticos AAT, FC e DBac, respectivamente, que apresentam assimetria.

Na Tabela 7 estão apresentados os critérios de convergência de Raftery e Lewis (1992) e Gelman e Rubin (1992). De acordo com os critérios de convergência as cadeias atingiram convergência, visto que os valores dos fatores de dependência (F_D), pelo critério de Raftery e Lewis, foram menores que 5. O critério de Gelman e Rubin também sugere convergência das cadeias, uma vez que os valores dos fatores de redução (\hat{R}) foram iguais a 1. Os resultados da análise realizada por meio dos critérios apresentados na Tabela 7, estão de acordo com a análise gráfica realizada anteriormente.

Tabela 7: Critérios de convergência de Raftery e Lewis (F_D) e Gelman e Rubin (R).

Parâmetros	AAT		2-ME		FC		DBac	
	F_D	\hat{R}	F_D	\hat{R}	F_D	\hat{R}	F_D	\hat{R}
δ	0,997	1,000	0,991	1,000	1,006	1,000	1,000	1,000
S	1,004	1,000	1,010	1,000	0,998	1,000	1,007	1,000
E	1,000	1,000	1,026	1,000	1,022	1,000	1,013	1,000

Dentre os 435 animais com suspeita de brucelose bovina, 226 (51,90%), 358 (82,30%), 363 (83,50%) e 173 (39,80%), foram classificados como positivos pelos testes AAT, 2-ME, FC e DBac, respectivamente. Os correspondentes testes classificaram como negativos, 209 (48,06%), 77 (17,70%), 72 (16,55%) e 262 (60,23%) animais.

Na Tabela 8 apresenta-se as médias das distribuições a posteriori dos parâmetros S , E e δ para os testes AAT, 2-ME, FC e DBac, bem como os respectivos desvios padrão (DP) e intervalos com 95% de credibilidade (ICr). Dentre os quatro testes diagnósticos, o AAT apresentou maior sensibilidade com aproximadamente 97%, enquanto os testes FC, 2-ME e DBac próximos de 96%, 93% e 54%, respectivamente. Todos os testes forneceram especificidades acima de 90%, com 99% para os testes AAT, FC e DBac, enquanto o teste 2-ME apresentou aproximadamente 91%. As estimativas das especificidades (E_3 e E_4) dos testes FC e DBac, respectivamente, foram semelhantes, uma vez que seus intervalos com 95% de credibilidade se sobrepõem. A sensibilidade e especificidade do teste 2-ME foram relativamente semelhantes às estimativas apresentadas por Souza (2013). O autor, utilizando o modelo de classe latente

Bayesiano, avaliou individualmente os testes diagnósticos AAT, 2-ME e DBac, bem como combinados, considerando os testes condicionalmente independentes. Na avaliação individual, o autor obteve os seguintes resultados para o teste 2-ME: $S_2 = 0,933$; $ICr_{95\%} = (0,900; 0,961)$ e $E_2 = 0,911$; $ICr_{95\%} = (0,884; 935)$.

Tabela 8: Médias das distribuições a posteriori dos parâmetros de desempenho dos testes AAT, 2-ME, FC e DBac e da prevalência, com respectivos desvios padrão (DP) e intervalos com 95% de credibilidade (ICr).

Testes	Parâmetros	Média	DP	ICr
AAT	δ	0,583	0,021	(0,542; 0,625)
	S_1	0,972	0,009	(0,949; 0,988)
	E_1	0,998	0,001	(0,997; 0,999)
2-ME	δ	0,804	0,019	(0,765; 0,841)
	S_2	0,934	0,015	(0,903; 0,961)
	E_2	0,911	0,013	(0,884; 0,935)
FC	δ	0,813	0,017	(0,778; 0,846)
	S_3	0,963	0,009	(0,943; 0,978)
	E_3	0,995	0,003	(0,988; 0,998)
DBac	δ	0,714	0,031	(0,652; 0,773)
	S_4	0,539	0,032	(0,479; 0,603)
	E_4	0,995	0,003	(0,988; 0,998)

Para o teste AAT, verifica-se que a estimativa da especificidade, 99,80%, foi superior a estimativa da sensibilidade, 97,20%. Por ser de triagem, espera-se que o teste AAT apresente sensibilidade maior que a especificidade (MEIRELLES-BARTOLI; MATHIAS, 2010). No entanto, resultados inversos podem ocorrer, como verificado em: Matope et al. (2011), que obteve $S_1 = 0,947$; $ICr_{95\%} = (0,874; 0,989)$ e $E_1 = 0,990$; $ICr_{95\%} = (0,980; 0,996)$.

As estimativas dos parâmetros de desempenho do teste DBac estão de acordo com as propriedades do teste, isto é, o diagnóstico bacteriológico tem como característica a alta especificidade (MOL et al., 2012; MANISH et al., 2013) e baixa sensibilidade (ABUBAKAR; MANSOOR; ARSHED, 2012). O testes 2-ME e FC são confirmatórios, desse modo apresentam como propriedade maior especificidade (BRASIL, 2006; MEIRELLES-BARTOLI; MATHIAS, 2010), no entanto, verifica-se que apenas o teste FC apresenta concordância com suas características, isto é, maior especificidade e menor sensibilidade. Conforme Mol et al. (2012), esse teste apresenta sensibilidade e especificidade variando de 77,1% a 100% e 65,0% a 100%, respectivamente. Embora a estimativa da sensibilidade tenha sido superior para o teste 2-ME, ainda conforme Mol et al. (2012), esse teste possui sensibilidade variando entre 88,4% e 99,6% e especificidade entre 91,5% e 99,8%, o que está de acordo com os resultados obtidos.

Ainda conforme a Tabela 8, observa-se que os desvios padrão para as sensibilidades

foram superiores quando comparados com os desvios padrão das especificidades, indicando que as estimativas das especificidades foram mais precisas que as estimativas das sensibilidades. Os parâmetros que apresentaram intervalos de credibilidade com maior amplitude foram: a prevalência (δ) e a sensibilidade (S_4) do teste DBac, com 0,121 e 0,124, respectivamente. Aqueles com menores amplitudes foram para as especificidades (E_1 , E_3 e E_4) dos testes AAT, FC e DBac, respectivamente, com 0,002, 0,01 e 0,01. Os demais parâmetros apresentaram intervalos com amplitude variando entre 0,035 e 0,083.

Os gráficos das distribuições a priori e a posteriori para a prevalência, sensibilidades e especificidades dos testes AAT, 2-ME, FC e DBac estão apresentados nas Figuras 5 a 7, respectivamente. Na Figura 7 observa-se que as distribuições a posteriori das especificidades se sobrepõem ou estão relativamente próximas às respectivas distribuições a priori. Acredita-se que tal resultado está relacionado a pouca informação fornecida pelos dados. Nesse caso, pode-se verificar a sobreposição ou proximidade entre as distribuições a posteriori (ICr) das especificidades, Tabela 8, e suas respectivas informações a priori, Tabela 6, seção 3.3, isto é, dos intervalos $[0,998; 0,999]$, $[0,890; 0,940]$, $[0,990; 1,00]$ e $[0,990; 1,00]$ para os testes AAT, 2-ME, FC e DBac. Tal comportamento também foi verificado para a prevalência do teste DBac, Figura 5, sua informação a priori foi $[0,630; 0,770]$.

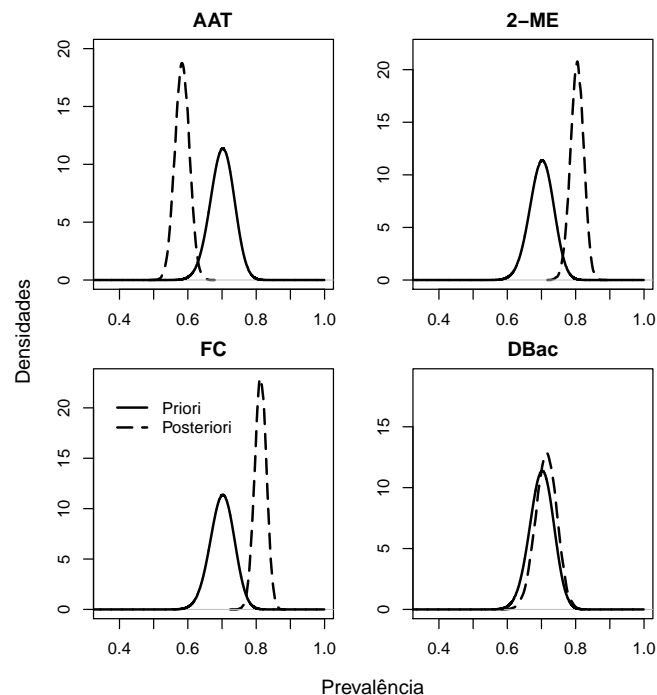


Figura 5: Gráficos das densidades a posteriori e priori da prevalência da doença (δ) para os testes AAT, 2-ME, FC e DBac.

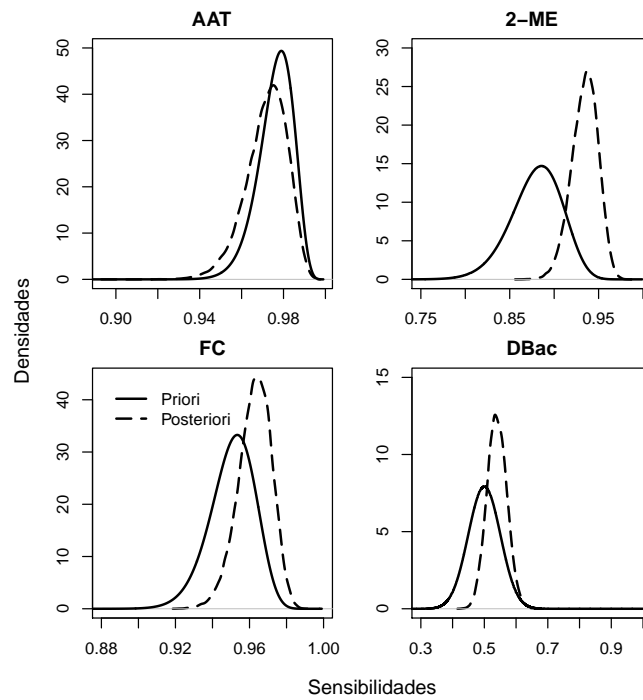


Figura 6: Gráficos das densidades a posteriori e priori das sensibilidades (S) para os testes AAT, 2-ME, FC e DBac.

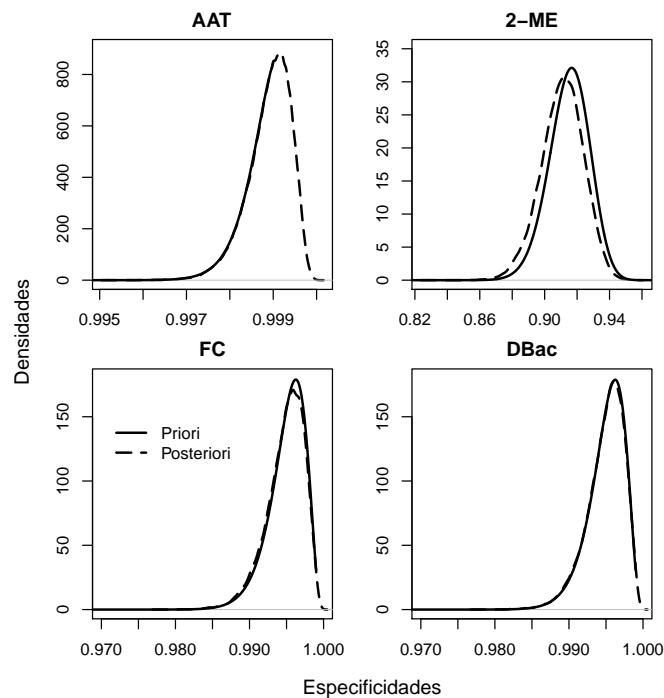


Figura 7: Gráficos das densidades a posteriori e priori das especificidades (E) para os testes AAT, 2-ME, FC e DBac.

Conforme já relatado neste texto, não existe um teste Padrão Ouro para diagnosticar a brucelose bovina, isto é, um teste diagnóstico que apresente sensibilidade e especificidade iguais a 100%. Uma alternativa para contornar essa limitação é combinar três ou mais testes diagnósticos imperfeitos para estimar os parâmetros de desempenho dos testes e a prevalên-

cia da doença (JOSEPH; GYORKOS; COUPAL, 1995; SANOGO; ABATIH; SAEGERMAN, 2014). Sendo os testes diagnósticos AAT, 2-ME, FC e DBac imperfeitos, pois apresentam sensibilidades e especificidades menores que 100% conforme declara Speybroeck et al. (2013), ao combiná-los surge uma importante questão investigada no presente trabalho: os testes podem ser condicionalmente dependentes, conforme descrito na seção 2.3.1, sendo a dependência condicional incorporada no modelo por meio de covariâncias condicionais entre os testes diagnósticos dado o estado da doença. Uma vez calculadas as covariâncias condicionais, foram obtidas as correlações condicionais. Os resultados das análises para os testes diagnósticos AAT, 2-ME, FC e DBac combinados, considerando-se independência e dependência condicional entre os testes, estão apresentados a seguir.

Na Tabela 9 estão apresentados os fatores de dependência (F_D) e de redução (R) fornecidos pelos critérios de Raftery e Lewis (1992) e Gelman e Rubin (1992), respectivamente, para os parâmetros: prevalência, sensibilidades, especificidades e correlações condicionais dos modelos. O modelo considerando testes condicionalmente independentes corresponde ao Modelo 1. Nos modelos considerando testes condicionalmente dependentes foram incluídas as covariâncias condicionais (dados não apresentados). E também as correlações condicionais ρ_D e ρ_{ND} para avaliar possíveis dependências entre os resultados dos testes diagnósticos: AAT e 2-ME (Modelo 2), AAT e FC (Modelo 3) e 2-ME e FC (Modelo 4), para os animais doentes (D) e não doentes (ND). O modelo considerando todos os pares de covariâncias condicionais e correlações condicionais citados acima, denominamos de Modelo 5.

Tabela 9: Critérios de convergência de Raftery e Lewis (F_D) e Gelman e Rubin (\hat{R}).

Parâmetros	Modelo 1		Modelo 2		Modelo 3		Modelo 4		Modelo 5	
	F_D	\hat{R}	F_D	\hat{R}	F_D	\hat{R}	F_D	\hat{R}	F_D	\hat{R}
δ	1,000	1,000	1,020	1,000	1,010	1,000	0,995	1,000	0,995	1,000
S_1	1,030	1,000	1,020	1,000	1,010	1,000	1,010	1,000	1,020	1,000
E_1	1,160	1,000	1,120	1,000	1,080	1,000	1,190	1,000	1,000	1,000
S_2	0,997	1,010	0,997	1,000	1,010	1,000	0,998	1,000	1,010	1,000
E_2	1,010	1,000	1,030	1,000	1,010	1,000	1,000	1,000	1,010	1,000
S_3	1,000	1,000	1,010	1,000	1,030	1,000	1,020	1,000	1,010	1,000
E_3	1,060	1,000	1,040	1,000	1,010	1,000	1,020	1,000	1,010	1,000
S_4	0,998	1,010	0,997	1,000	1,010	1,000	1,020	1,000	0,999	1,000
E_4	1,040	1,000	1,030	1,000	1,010	1,000	1,050	1,000	1,020	1,000
$\rho_{D_{12}}$			0,990	1,000					1,010	1,000
$\rho_{ND_{12}}$			0,993	1,000					1,000	1,000
$\rho_{D_{13}}$					0,997	1,000			1,010	1,000
$\rho_{ND_{13}}$					1,000	1,000			0,988	1,000
$\rho_{D_{23}}$							1,030	1,000	1,010	1,000
$\rho_{ND_{23}}$							1,000	1,000	1,000	1,000

Assim como os resultados obtidos tomando cada teste individualmente, verifica-se na análise conduzida para os testes combinados que as cadeias alcançaram convergência, pois os valores dos fatores de dependência (F_D), foram menores que 5 e, os valores dos fatores de redução (\hat{R}), foram iguais a 1. Os gráficos dos traços, das autocorrelações e densidades para os parâmetros dos Modelos 1 e 2 estão apresentados no Apêndice A, para os demais modelos, resultados similares foram encontrados.

As médias das distribuições a posteriori dos parâmetros: prevalência (δ), sensibilidades (S), especificidades (E) e correlações condicionais (ρ_D e ρ_{ND}) para os Modelos 1, 2, 3, 4 e 5, com respectivos intervalos com 95% de credibilidade (ICr), bem como o critério de informação da deviance (DIC) e o número efetivo de parâmetros (p_D) dos modelos, estão apresentados na Tabela 10.

Na Tabela 10 observa-se que apenas os intervalos com 95% de credibilidade dos parâmetros das correlações condicionais entre os testes: AAT e 2-ME no Modelo 2 ($\rho_{D_{12}}$) e AAT e FC nos Modelos 3 e 5 ($\rho_{ND_{13}}$) não incluem o zero. No entanto, a correlação condicional entre os testes AAT e 2-ME no Modelo 2 foi de $\rho_{D_{12}} = 0,103$. Quando as correlações condicionais são menores ou iguais a 0,20, os modelos condicionalmente independentes e condicionalmente dependentes apresentam estimativas semelhantes (GEORGIADIS et al., 2003; RAHMAN et al., 2013). A média da distribuição a posteriori da correlação condicional entre os testes AAT e FC para os animais não portadores da doença foram iguais a $\rho_{ND_{13}} = 0,223$; $\text{ICr}_{95\%} = (0,007; 0,610)$, no Modelo 3, e $\rho_{ND_{13}} = 0,227$; $\text{ICr}_{95\%} = (0,007; 0,622)$ no Modelo 5. Nota-se que as correlações condicionais nos Modelos 3 e 5 não provocaram diferenças entre as médias a posteriori dos parâmetros dos modelos, quando comparadas com as as médias a posteriori do Modelo 1, uma vez que os intervalos com 95% de credibilidade estão sobrepostos. De acordo com Praud et al. (2012), quando dois intervalos de credibilidade se sobrepõem, os parâmetros não podem ser considerados diferentes. Acredita-se que esse resultado pode ser devido a elevada amplitude apresentada pelos intervalos de credibilidade, com o limite inferior próximo de zero.

Tabela 10: Médias das distribuições a posteriori da prevalência, sensibilidades, especificidades e correlações condicionais dos modelos 1, 2, 3, 4 e 5, com respectivos intervalos com 95% de credibilidade (ICr), critério de informação da deviance (DIC) e o número efetivo de parâmetros (p_D).

Parâmetros	Modelo 1		Modelo 2		Modelo 3		Modelo 4		Modelo 5	
	Média	ICr	Média	ICr	Média	ICr	Média	ICr	Média	ICr
δ	0,890	(0,863; 0,915)	0,890	(0,863; 0,915)	0,888	(0,861; 0,914)	0,890	(0,863; 0,916)	0,885	(0,856; 0,912)
S_1	0,729	(0,696; 0,760)	0,729	(0,696; 0,761)	0,728	(0,695; 0,759)	0,728	(0,696; 0,760)	0,728	(0,695; 0,760)
E_1	0,998	(0,997; 0,999)	0,998	(0,998; 0,999)	0,999	(0,998; 0,999)	0,998	(0,997; 0,999)	0,998	(0,997; 0,999)
S_2	0,846	(0,815; 0,875)	0,857	(0,826; 0,886)	0,846	(0,814; 0,875)	0,844	(0,812; 0,873)	0,852	(0,819; 0,882)
E_2	0,901	(0,873; 0,927)	0,901	(0,872; 0,926)	0,899	(0,871; 0,926)	0,904	(0,876; 0,929)	0,899	(0,870; 0,925)
S_3	0,901	(0,877; 0,922)	0,901	(0,877; 0,922)	0,899	(0,875; 0,921)	0,900	(0,877; 0,922)	0,896	(0,871; 0,918)
E_3	0,994	(0,987; 0,998)	0,994	(0,988; 0,999)	0,994	(0,987; 0,999)	0,995	(0,988; 0,999)	0,994	(0,988; 0,999)
S_4	0,428	(0,386; 0,471)	0,428	(0,385; 0,471)	0,428	(0,386; 0,472)	0,428	(0,385; 0,471)	0,427	(0,384; 0,470)
E_4	0,994	(0,988; 0,999)	0,994	(0,988; 0,999)	0,994	(0,988; 0,998)	0,994	(0,988; 0,998)	0,992	(0,983; 0,998)
$\rho_{D_{12}}$			0,103	(0,016; 0,193)					0,022	(-0,007; 0,052)
$\rho_{ND_{12}}$			0,042	(-0,008; 0,113)					0,042	(-0,008; 0,112)
$\rho_{D_{13}}$					-0,034	(-0,107; 0,046)			-0,033	(-0,064; -0,045)
$\rho_{ND_{13}}$					0,223	(0,007; 0,610)			0,227	(0,007; 0,622)
$\rho_{D_{23}}$							-0,021	(-0,089; 0,063)	0,018	(-0,051; 0,018)
$\rho_{ND_{23}}$							0,084	(-0,019; 0,243)	0,074	(-0,019; 0,249)
DIC		220,2		218,0		220,9		221,4		218,9
p_D		3,443		4,435		4,419		4,375		5,963

Com relação aos valores dos DICs, observa-se que o menor valor foi apresentado pelo Modelo 2. Comparando o DIC do Modelo 2 com os demais, verifica-se que todas as diferenças foram menores que cinco. Nesse caso, como descrito na seção 3.4 e conforme os resultados apresentados na Tabela 10, há evidências que o ajuste dos Modelos 1, 2, 3, 4 e 5 aos dados são semelhantes, pois as médias das distribuições a posteriori dos parâmetros não diferem entre os modelos. Nessa situação, como já mencionado, a seleção entre os modelos é baseada no número efetivo de parâmetros, p_D . Sendo o menor valor p_D observado para o Modelo 1, com $p_D = 3,443$, tem-se que o Modelo 1 é o melhor para descrever o conjunto de dados em estudo. Também pode-se concluir que a pressuposição de independência condicional entre os testes AAT, 2-ME e FC é satisfeita para o conjunto de dados em avaliação.

De acordo com os resultados do Modelo 1, assim como nos testes avaliados individualmente, as estimativas das especificidades foram mais precisas que as estimativas das sensibilidades, conforme os resultados dos desvios padrão (Apêndice B). A sensibilidade (S_4) do teste DBac foi o parâmetro que apresentou intervalo de credibilidade com maior amplitude, com 0,085. Os demais parâmetros apresentaram intervalos com amplitude variando entre 0,002 e 0,064.

Com base na Tabela 10, verifica-se uma alta estimativa para a prevalência da doença, sendo $\delta = 0,890$. Acredita-se que esse valor está de acordo com as características do conjunto de dados em estudo, pois este é constituído apenas de amostras de sangue ou carcaças de animais com suspeita de brucelose bovina, provenientes do Serviço de Inspeção Federal (SIF), e encaminhadas ao laboratório para confirmação da suspeita.

Dentre os testes, o FC apresentou a maior sensibilidade com aproximadamente 90%, enquanto os testes AAT, 2-ME e DBac com sensibilidade próxima de 73%, 85% e 43%, respectivamente. Com relação as especificidades, observa-se que todas foram acima de 90%, com 99% para os testes AAT, FC e DBac, enquanto o teste 2-ME apresentou 90,1%.

Como já referido nesse texto, os testes 2-ME, FC e DBac têm como característica maior especificidade e menor sensibilidade. Conforme os resultados da Tabela 10, verifica-se que esses testes estão em concordância com suas propriedades. O teste AAT, assim como no caso dos testes avaliados individualmente, apresentou maior especificidade e menor sensibilidade. A sensibilidade do teste AAT, no presente estudo, foi relativamente maior que a estimativas encontradas por Sanogo et al. (2013), com $S_1 = 0,547$; $ICr_{95\%} = (0,235; 0,950)$ e, Ahasan et al. (2017), com $S_1 = 0,578$; $ICr_{95\%} = (0,330; 0,790)$. Os autores utilizaram o modelo

de classe latente Bayesiano para estimar os parâmetros de desempenho do teste AAT, aplicado para diagnosticar a brucelose bovina em animais: da Costa do Marfim (SANOGO et al., 2013) e Bangladeche (AHASAN et al., 2017).

A sensibilidade e especificidade do teste AAT foram semelhantes as estimativas obtidas por Arifa et al. (2018), com $S_1 = 0,716$; $ICr_{95\%} = (0,575; 0,852)$ e $E_1 = 0,993$; $ICr_{95\%} = (0,975; 0,999)$, respectivamente. Arifa et al. (2018), estimaram a sensibilidade e especificidade do teste AAT, aplicado para diagnosticar a brucelose bovina em animais do Paquistão, empregando o modelo de classe latente Bayesiano.

O teste AAT é bastante sensível (JARDIM et al., 2009), conforme Corbel (2006), apresenta sensibilidade superior a 99%. No entanto, alguns autores têm relatado baixa estimativa para a sensibilidade desse teste (SANOGO et al., 2013; AHASAN et al., 2017; ARIFA et al., 2018). Segundo os autores, esse fato pode ocorrer devido aos resultados falsos negativos, que acontecem particularmente no estágio inicial da infecção aguda (CORBEL, 2006), em casos de longa evolução da doença (crônica) (DÍAZ et al., 2011) e/ou devido ao fenômeno prozona (ARIFA et al., 2018). Nesse fenômeno ocorrem resultados falsos negativos devido ao excesso de anticorpos que interfere na formação de aglutinação necessária para visualizar os resultados do teste (DÍAZ et al., 2011; ARIFA et al., 2018). Devido a vacina B19, utilizada em todo território nacional por recomendações do PNCEBT, resultados falsos positivos também podem ocorrer (BRASIL, 2006), nesse caso, recomenda-se confirmar os resultados por meio de um ou mais testes diagnósticos com alta especificidade (POESTER et al., 2010), por exemplo, os testes 2-ME, FC e DBac. Esse procedimento é adotado para evitar o sacrifício de animais não doentes (BRASIL, 2006) e, conseqüentemente, perdas econômicas.

Embora o AAT, em alguns estudos, não apresente as características esperadas para uma teste de triagem (isto é, $S > E$), por ser um teste de baixo custo e de fácil execução (JARDIM et al., 2009; BRASIL, 2006), é mundialmente aplicado para o diagnóstico da brucelose bovina em programas de combate à doença (SANOGO et al., 2013).

5 CONCLUSÕES

No presente estudo, o modelo de classe latente mostrou-se adequado para estimar a sensibilidade e especificidade dos testes diagnósticos AAT, 2-ME, FC e DBac, bem como a prevalência da doença, uma vez que o verdadeiro estado doença é desconhecido. O enfoque Bayesiano também mostrou-se adequado no estudo do desempenho dos testes diagnósticos, pois todos os parâmetros de interesse foram estimados, algo que no enfoque Frequentista tende a ser de difícil execução à medida que aumenta-se o número de testes e, conseqüentemente, o número de parâmetros.

Ao avaliar os quatro testes individualmente, verificou-se que nenhum dos quatro testes pode ser utilizado sozinho para o diagnosticar a brucelose bovina. Nos resultados da análise dos quatro testes diagnósticos combinados, observou-se que os testes AAT, 2-ME e FC são condicionalmente independentes, no entanto, no estudo do desempenho de testes diagnósticos, quando há suspeita de dependência condicional entre os testes, esta não deve ser ignorada, pois evita-se possíveis estimativas tendenciosas para a sensibilidade e a especificidade e, conseqüentemente, tomadas de decisões incorretas pelo pesquisador.

A sensibilidade do teste AAT foi relativamente baixa, quando comparada com alguns autores. Devido à sua importância em programas de combate à brucelose, como o PNCEBT, recomendamos que constantes estudos sejam realizados sobre o seu desempenho.

6 REFERÊNCIAS BIBLIOGRÁFICAS

- ABUBAKAR, M.; MANSOOR, M.; ARSHED, M. J. Bovine brucellosis: Old and new concepts with pakistan perspective. **Pakistan Veterinary Journal**, v. 32, n. 2, p. 147–155, 2012.
- ADEL, A.; BERKVEN, D. Modelling conditional dependence between multiple diagnostic tests, using co-variances between test results. **Antwerp**, p. 1 – 6, 2002.
- AHASAN, M. S. et al. Bovine and caprine brucellosis in Bangladesh: Bayesian evaluation of four serological tests, true prevalence, and associated risk factors in household animals. **Tropical animal health and production**, v. 49, n. 1, p. 1 – 11, 2017.
- ALBERT, P. S.; DODD, L. E. A cautionary note on the robustness of latent class models for estimating diagnostic error without a gold standard. **Biometrics**, v. 60, p. 427 – 435, June 2004.
- ALLMAN, E. S.; MATIAS, C.; RHODES, J. A. Identifiability of parameters in latent structure models with many observed variables. **The Annals of Statistics**, v. 37, n. 6 A, p. 3099 – 3132, 2009.
- ALVES, A. et al. Economic analysis of vaccination to control bovine brucellosis in the states of são paulo and mato grosso, brazil. **Preventive Veterinary Medicine**, v. 118, p. 351 – 358, 2015.
- ARAGON, D. C.; MARTINEZ, E. Z.; ACHCAR, J. A. Bayesian estimation for performance measures of two diagnostic tests in the presence of verification bias. **Journal of Biopharmaceutical Statistics**, v. 20, p. 821 – 834, 2010.
- ARIFA, S. et al. Evaluation of three serological tests for diagnosis of bovine brucellosis in smallholder farms in pakistan by estimating sensitivity and specificity using bayesian latent class analysis. **Preventive Veterinary Medicine**, v. 149, p. 21 – 28, 2018.
- BASSI, F. Identification of latent class markov models with multiple indicators and correlated measurement errors. **Journal of the Italian Statistical Society**, v. 6, n. 3, p. 201 – 211, 1997.
- BERG, A.; MEYER, R.; YU, J. Deviance information criterion for comparing stochastic volatility models. **Journal of Business & Economic Statistics**, v. 22, n. 1, p. 107 – 120, 2004.
- BERGER, J. O. **Statistical Decision Theory and Bayesian Analysis**. 2. ed. New York: Springer series in statistics, 1985. 617 p.
- BEST, N. G.; COWLES, M. K.; VINES, K. Coda* convergence diagnosis and output analysis software for gibbs sampling output version 0.30. **MRC Biostatistics Unit**, Cambridge, v. 52, p. 1 – 42, 1995.
- BOLSTAD, W. M. **Understanding Computational Bayesian Statistics**. New Jersey: John Wiley & Sons, 2010. 320 p.
- BOTAIER NETO, M. et al. Brucelose bovina. **Revista Científica Eletrônica de Medicina Veterinária**, Ano VII, n. 12, p. 1 – 6, 2009.

BOX, G. E. P.; TIAO, G. C. **Bayesian Inference in Statistical Analysis**. Massachusetts: Addison-Wesley Publishing Company, 1973. 588 p.

BRANSCUM, A. J.; GARDNER, I. A.; JOHNSON, W. O. Estimation of diagnostic-test sensitivity and specificity through bayesian modeling. **Preventive Veterinary Medicine**, v. 68, n. 2 - 4, p. 145 – 163, 2005.

BRASIL. **Programa Nacional de Controle e Erradicação da Brucelose e da Tuberculose Animal (PNCEBT)**. Brasília: MAPA/DAS/DSA, 2006. 188 p. Disponível em: <http://www.agricultura.gov.br/arq_editor/file/Aniamal/programa%20nacional%20sanidade%20brucelose/Manual%20do%20PNCEBT%20-%20Original.pdf>. Acesso em: 26 fev. 2016.

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. **Relatório de Gestão do Exercício de 2010, LANAGRO-MG**. [S.l.], 2011. Disponível em: <<http://www.agricultura.gov.br/aceso-a-informacao/auditorias/2010/lanagro-mg.pdf/view>>. Acesso em: 07 dez. 2017.

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. **Lanagros**. 2017. Disponível em: <<http://www.agricultura.gov.br/assuntos/laboratorios/lanagros>>. Acesso em: 07 dez. 2017.

BROOKS, S. P. Markov chain monte carlo method and its application. **Journal of the Royal Statistical Society**, v. 47, n. 1, p. 69 – 100, 1998.

BROOKS, S. P.; GELMAN, A. General methods for monitoring convergence of iterative simulations. **Journal of Computational and Graphical Statistics**, v. 7, n. 4, p. 434 – 455, 1998.

BROOKS, S. P.; ROBERTS, G. O. Assessing convergence of markov chain monte carlo algorithms. **Statistics and Computing**, v. 8, n. 4, p. 319 – 335, 1998.

BROWN, E. R. Bayesian estimation of the time-varying sensitivity of a diagnostic test with application to mother-to-child transmission of HIV. **Biometrics**, v. 66, p. 1266 – 1274, 2010.

BROWNE, R. H. Using the sample range as a basis for calculating sample size in power calculations. **The American Statistician**, v. 55, n. 4, p. 293 – 298, 2001.

CARLIN, B. P.; LOUIS, T. A. **Bayesian Methods for Data Analysis**. 3. ed. [S.l.]: Chapman & Hall/CRC, 2008. 529 p.

CASELLA, G.; BERGER, R. L. **Inferência Estatística**. 2. ed. São Paulo: Cengage Learning, 2010. 588 p.

CHRISTENSEN, R. et al. **Bayesian Ideas and Data Analysis: An Introduction for Scientists and Statisticians**. Boca Raton, Flórida, USA: Taylor & Francis, 2011. 498 p.

CLEMENTINO, I. J.; AZEVEDO, S. S. de. Bovine brucellosis: epidemiological situation in Brazil and disease control initiatives. **Semina: Ciências Agrárias**, v. 37, n. 4, p. 2021 – 2034, julho/agosto 2016.

COLLINS, J.; HUYNH, M. Estimation of diagnostic test accuracy without full verification: a review of latent class methods. **Statistics in Medicine**, v. 33, n. 24, p. 4141 – 4169, october 2014.

COLLINS, L. M.; LANZA, S. T. **Latent Class and Latent Transition Analysis**. New Jersey: John Wiley & Sons, 2010. 285 p.

CONAB. Companhia Nacional de Abastecimento. **Leite e Derivados**. 2017. Disponível em: <http://www.conab.gov.br/OlalaCMS/uploads/arquivos/17\05\15\14\13\38\leite_abril\2017.pdf>. Acesso em: 09 dez. 2017.

CORBEL, M. **Brucellosis in Humans and Animals**. Genebra, Suíça: World Health Organisation, 2006. Disponível em: <<http://www.who.int/csr/resources/publications/Brucellosis.pdf>>. Acesso em: 01 set. 2015.

COWLES, M. K. **Applied Bayesian Statistics With R and OpenBUGS Examples**. London: Springer-Verlag New York, 2013. v. 98. 232 p. (Springer Texts in Statistics, v. 98).

COWLES, M. K.; CARLIN, B. P. Markov chain monte carlo convergence diagnostics: A comparative review. **Journal of the American Statistical Association**, v. 91, n. 434, p. 883 – 904, June 1996.

COWLES, M. K.; ROSENTHAL, J. S. A simulation approach to convergence rates for markov chain monte carlo algorithms. **Statistics and Computing**, v. 8, n. 2, p. 115 – 124, 1998.

DeGROOT, M. H.; SCHERVISH, M. J. **Probability and Statistics**. 4. ed. Boston: Pearson, 2012. 912 p.

DENDUKURI, N.; BÉLISLE, P.; JOSEPH, L. Bayesian sample size for diagnostic test studies in the absence of a gold standard: Comparing identifiable with non-identifiable models. **Statistics in Medicine**, v. 29, p. 2688 – 2697, 2010.

DENDUKURI, N.; HADGU, A.; WANG, L. Modeling conditional dependence between diagnostic tests: A multiple latent variable model. **Statistics in Medicine**, v. 28, p. 441 – 461, 2009.

DENDUKURI, N.; JOSEPH, L. Bayesian approaches to modeling the conditional dependence between multiple diagnostic tests. **Biometrics**, v. 57, p. 158 – 167, 2001.

DÍAZ, R. et al. The rose bengal test in human brucellosis: A neglected test for the diagnosis of a neglected disease. **PLOS Neglected tropical diseases**, v. 5, n. 4, p. 1 – 7, 2011.

ENØE, C.; GEORGIADIS, M. P.; JOHNSON, W. O. Estimation of sensitivity and specificity of diagnostic tests and disease prevalence when the true disease state is unknown. **Preventive Veterinary Medicine**, v. 45, p. 61 – 81, 2000.

FORMANN, A. K. Latent class model diagnosis from a frequentist point of view. **Biometrics**, v. 59, p. 189 – 196, march 2006.

GAMERMAN, D.; LOPES, H. F. **Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference**. 2. ed. [S.l.]: Chapman & Hall/CRC, 2006.

GARDNER, I. A. The utility of bayes theorem and bayesian inference in veterinary clinical practice and research. **Australian Veterinary Journal**, v. 80, n. 12, p. 758 – 761, 2002.

GARDNER, I. A.; HIETALA, S.; BOYCE, W. Validity of using serological tests for diagnosis of diseases in wild animals. **Revue Scientifique et Technique (International Office of Epizootics)**, v. 15, n. 1, p. 323 – 335, 1996.

- GARDNER, I. A. et al. Conditional dependence between tests affects the diagnosis and surveillance of animal diseases. **Preventive Veterinary Medicine**, v. 45, p. 107 – 122, 2000.
- GELMAN, A. et al. **Bayesian Data Analysis**. 2. ed. Boca Raton, Flórida, USA: Chapman & Hall/CRC Texts in Statistical Science, 2004. 668 p.
- GELMAN, A.; RUBIN, B. D. Inference from iterative simulation using multiple sequences. **Statistical Science**, v. 7, n. 4, p. 457 – 472, July 1992.
- GEORGIADIS, M. P. et al. Correlation-adjusted estimation of sensitivity and specificity of two diagnostic tests. **Journal of the Royal Statistical Society: Series C (Applied Statistics)**, v. 52, n. 1, p. 63 – 76, 2003.
- GERESU, M. A.; KASSA, G. M. A review on diagnostic methods of brucellosis. **Veterinary Science & Technology**, v. 7, n. 323, p. 1 – 8, 2016.
- GETACHEW, T. et al. Bayesian estimation of sensitivity and specificity of rose bengal, complement fixation, and indirect elisa tests for the diagnosis of bovine brucellosis in ethiopia. **Veterinary medicine international**, v. 2016, p. 1 – 5, 2016.
- GODFROID, J. et al. Brucellosis at the animal/ecosystem/human interface at the beginning of the 21st century. **Preventive Veterinary Medicine**, v. 102, p. 118 – 131, 2011.
- GOODMAN, L. A. Exploratory latent structure analysis using both identifiable and unidentifiable models. **Biometrika**, v. 51, n. 2, p. 215 – 231, 1974.
- GORSICH, E. E. et al. Evaluation of the sensitivity and specificity of an enzyme-linked immunosorbent assay for diagnosing brucellosis in african buffalo (*syncerus caffer*). **Journal of Wildlife Diseases**, v. 51, p. 9 – 18, 2015.
- GREINER, M.; GARDNER, I. A. Application of diagnostic tests in veterinary epidemiologic studies. **Preventive Veterinary Medicine**, v. 45, p. 43 – 59, 2000a.
- GREINER, M.; GARDNER, I. A. Epidemiologic issues in the validation of veterinary diagnostic tests. **Preventive Veterinary Medicine**, v. 45, p. 3 – 22, 2000b.
- HADGU, A.; DENDUKURI, N.; HILDEN, J. Evaluation of nucleic acid amplification tests in the absence of a perfect gold-standard test: A review of the statistical and epidemiologic issues. **Epidemiology**, v. 16, n. 5, p. 604 – 612, 2005.
- HADGU, A.; QU, Y. A biomedical application of latent class models with random effects. **Journal of the Royal Statistical Society**, v. 47, n. 4, p. 603 – 616, 1998.
- HAMRA, G.; MACLEHOSE, R.; RICHARDSON, D. Markov chain monte carlo: an introduction for epidemiologists. **International journal of epidemiology**, v. 42, n. 2, p. 627 – 634, 2013.
- HANSON, T.; JOHNSON, W. O.; GARDNER, I. A. Hierarchical models for estimating herd prevalence and test accuracy in the absence of a gold standard. **Journal of Agricultural, Biological, and Environmental Statistics**, v. 8, n. 2, p. 223 – 239, 2003.
- HUANG, G. H. Model identifiability. **Wiley StatsRef: Statistics Reference Online**, v. 3, p. 1249 – 1251, 2005.

- HUI, S. L.; WALTER, S. D. Estimating the error rates of diagnostic tests. **Biometrics**, v. 36, n. 1, p. 167 – 171, march 1980.
- HUI, S. L.; ZHOU, X. H. Evaluation of diagnostic tests without gold standards. **Statistical Methods in Medical Research**, v. 7, p. 354 – 370, 1998.
- JARDIM, G. C. et al. Comparação do elisa indireto no diagnóstico da brucelose em rebanho bovino vacinado e não vacinado. **Agrarian**, v. 2, n. 5, p. 131 – 142, jul./set. 2009.
- JONES, G. et al. Identifiability of models for multiple diagnostic testing in the absence of a gold standard. **Biometrics**, v. 66, n. 3, p. 855 – 863, september 2010.
- JOSEPH, L.; GYORKOS, T. W.; COUPAL, L. Bayesian estimation of disease prevalence and the parameters of diagnostic test in the absence of a Gold Standard. **American Journal of Epidemiology**, v. 141, n. 3, p. 263 – 272, 1995.
- KAPLAN, D. **Bayesian statistics for the social sciences**. 1. ed. New York: Guilford Publications, 2014. 318 p.
- KIST, B. B. et al. **Anuário brasileiro da pecuária 2017**. Santa Cruz do sul, RS, 2017. 59 p. Disponível em: <<http://www.editoragazeta.com.br/wp-content/uploads/2017/09/PDF-Pecu%C3%A1ria-2017.pdf>>. Acesso em: 07 dez. 2017.
- LAGE, A. P. et al. Brucelose bovina: uma atualização. **Revista Brasileira de Reprodução Animal**, v. 32, n. 3, p. 202 – 212, 2008.
- LANGHEHEINE, R. New developments in latent class theory. In: LANGHEHEINE, R.; ROST, J. (Ed.). **Latent Trait and Latent Class Models**. New York: Plenum Press, 1988. cap. 4, p. 77 – 108.
- LANZA, S. T. et al. Proc lca: A sas procedure for latent class analysis. **Structural Equation Modeling**, v. 14, n. 4, p. 671 – 694, 2007.
- LEE, S. **Structural Equation Modeling: A Bayesian Approach**. New York: John Wiley & Sons, 2007. 458 p. (Wiley Series in Probability and Statistics).
- LESAFFRE, E.; SPEYBROECK, N.; BERKVEN, D. Bayes and diagnostic testing. **Veterinary Parasitology**, v. 148, n. 1, p. 58 – 61, 2007.
- LUNN, D. et al. **The BUGS Book: A Practical Introduction to Bayesian Analysis**. [S.l.]: Taylor & Francis, 2012. 399 p. (Chapman & Hall/CRC Texts in Statistical Science).
- MALHEIROS, L. A. **Fim do embargo à carne bovina brasileira nos EUA**. 2016. Disponível em: <<http://www.olhardireto.com.br/artigos/exibir.asp?artigo=fim-do-embargo-a-carne-bovina-brasileira-nos-eua&id=7786>>. Acesso em: 26 nov. 2016.
- MANISH, K. et al. Brucellosis: An updated review of the disease. **Indian Journal of Animal Sciences**, v. 83, n. 1, p. 3 – 16, 2013.
- MARTINEZ, E. Z.; ACHCAR, J. A.; LOUZADA-NETO, F. Estimators of sensitivity and specificity in the presence of verification bias: A bayesian approach. **Computational Statistics & Data Analysis**, v. 51, p. 601 – 611, 2006.

- MARTINEZ, E. Z.; LOUZADA-NETO, F.; PEREIRA, B. B. de. A curva roc para testes diagnósticos. **Cadernos de Saúde Pública**, v. 11, n. 1, p. 7 – 31, 2003.
- MATOPE, G. et al. Evaluation of sensitivity and specificity of rbt, c-elisa and fluorescence polarisation assay for diagnosis of brucellosis in cattle using latent class analysis. **Veterinary Immunology and Immunopathology**, v. 141, p. 58 – 63, 2011.
- McCUTCHEON, A. L. **Latent Class Analysis**. California: Sage Publications, 1987. 96 p. (Quantitative Applications in the Social Sciences).
- McHUGH, R. B. Efficient estimation and local identification in latent class analysis. **Psychometrika**, v. 21, n. 4, p. 331 – 347, december 1956.
- MEIRELLES-BARTOLI, R. B.; MATHIAS, L. A. Estudo comparativo entre os testes adotados pelo pncebt para o diagnóstico sorológico da brucelose em bovinos. **Arquivos do Instituto Biológico**, v. 77, n. 1, p. 11 – 17, 2010.
- MENTEN, J.; BOELAERT, M.; LESAFFRE, E. Bayesian latent class models with conditionally dependent diagnostic tests: A case study. **Statistics in Medicine**, v. 27, p. 4469 – 4488, 2008.
- MOL, J. P. da S. et al. Diagnóstico laboratorial da brucelose. **Revista Brasileira de Ciência Veterinária**, v. 19, n. 3, p. 117 – 126, set./dez. 2012.
- NEELY-BARNES, S. Latent class models in social work. **Social Work Research**, v. 32, n. 2, p. 114 – 121, june 2010.
- NTZOUFRAS, I. **Bayesian modeling using WinBUGS**. [S.l.]: John Wiley & Sons, 2009. 492 p.
- O'NEILL, D. Measuring obesity in the absence of a gold standard. **Economics and Human Biology**, v. 17, p. 116 – 128, 2015.
- PAN-NGUM, W. et al. Estimating the true accuracy of diagnostic tests for dengue infection using bayesian latent class models. **PLoS ONE**, v. 8, n. 1, p. 1 – 7, 2013.
- PAULA, C. L. de et al. Detecção de brucella spp. em leite bovino não pasteurizado através da reação de cadeia pela polimerase (PCR). **Arquivos do Instituto Biológico**, v. 82, p. 1–5, 2015.
- PAULINO, C. D.; TURKMAN, M. A. A.; MURTEIRA, B. **Estatística Bayesiana**. Lisboa: Fundação Calouste Gulbenkian, 2003. 446 p.
- PEREIRA, G. de A. et al. A general latent class model for performance evaluation of diagnostic tests in the absence of a gold standard: an application to chagas disease. **Computational and Mathematical Methods in Medicine**, v. 2012, p. 1 – 12, 2012.
- POESTER, F. et al. Estudos de prevalência da brucelose bovina no âmbito do programa nacional de controle e erradicação de brucelose e tuberculose: Introdução. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v. 61, n. 1, p. 1 – 5, 2009.
- POESTER, F. P. et al. Diagnosis of brucellosis. **The Open Veterinary Science Journal**, v. 4, p. 46 – 60, 2010.

- PRAUD, A. et al. Evaluation of three competitive elisas and a fluorescence polarisation assay for the diagnosis of bovine brucellosis. **The Veterinary Journal**, v. 216, p. 38 – 44, 2016.
- PRAUD, A. et al. Estimation of sensitivity and specificity of five serological test for the diagnostic of porcine brucellosis. **Preventive Veterinary Medicine**, v. 104, n. 1-2, p. 94 – 100, 2012.
- PRESS, S. J. **Subjective and objective Bayesian statistics: principles, models, and applications**. 2. ed. New Jersey: John Wiley & Sons, 2003. 600 p.
- QU, Y.; TAN, M.; KUTNER, M. H. Random effects models in latent class analysis for evaluating accuracy of diagnostic tests. **Biometrics**, v. 52, p. 797 – 810, september 1996.
- R Core Team. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2016. Disponível em: <URLhttps://www.R-project.org/>. Acesso em: 02 mai. 2016.
- RAFTERY, A. E.; LEWIS, S. How many iterations in the gibbs sampler? In: BERNARDO, J. M. et al. (Ed.). **Bayesian Statistics 4**. Oxford: University Press, 1992. p. 763 – 773.
- RAHMAN, A. A. et al. Bayesian estimation of true prevalence, sensitivity and specificity of indirect elisa, rose bengal test and slow agglutination test for the diagnosis of brucellosis in sheep and goats in bangladesh. **Preventive Veterinary Medicine**, v. 110, p. 242 – 252, 2013.
- REITSMA, J. B. et al. A review of solutions for diagnostic accuracy studies with an imperfect or missing reference standard. **Journal of Clinical Epidemiology**, v. 62, p. 797 – 806, 2009.
- RHIEL, G. S.; MARKOWSKI, E. An update on using the range to estimate p when determining sample sizes. **Psychological Reports**, v. 120, n. 2, p. 319 – 331, 2017.
- RINDSKOPF, D.; RINDSKOPF, W. The value of latent class analysis in medical diagnosis. **Statistics in Medicine**, v. 5, n. 1, p. 21 – 27, 1986.
- ROBERT, C.; CASELLA, G. **Introducing Monte Carlo Methods with R**. New York: Springer, 2010. 284 p. (Use R!).
- ROBERT, C. P. **The bayesian choice: From Decision-theoretic foundations to computational implementation**. 2. ed. New York: Springer, 2007. 602 p.
- ROBINSON, A. **Guidelines for coordinated human and animal brucellosis surveillance**. Rome, 2003.
- SANOGO, M.; ABATI, E.; SAEGERMAN, C. Bayesian versus frequentist methods for estimating true prevalence of disease and diagnostic test performance. **The Veterinary Journal**, v. 202, p. 204 – 207, 2014.
- SANOGO, M. et al. Bayesian estimation of the true prevalence, sensitivity and specificity of the rose bengal and indirect elisa tests in the diagnosis of bovine brucellosis. **The Veterinary Journal**, v. 195, p. 114 – 120, 2013.
- SKRONDAL, A.; RABE-HESKETH, S. Latent variable modelling: A survey. **Scandinavian Journal of Statistics**, v. 34, n. 4, p. 712 – 745, 2007.

- SMEDEN, M. van et al. Latent class models in diagnostic studies when there is no reference standard - a systematic review. **American Journal of Epidemiology**, v. 179, n. 4, p. 423 – 431, november 2014.
- SOARES FILHO, P. M. et al. Comparação entre as provas de 2-mercaptoetanol e de fixação de complemento para o diagnóstico sorológico da brucelose bovina em animais com diferentes históricos vacinais. **Veterinária Notícias**, v. 17, n. 1, p. 26 – 32, 2011.
- SOLA, M. C. et al. Brucelose bovina: Revisão. **Enciclopédia Biósfera, Centro Científico Conhecer**, v. 10, n. 18, p. 686 – 714, 2014.
- SOUZA, M. R. S. **Estimação da sensibilidade e especificidade de testes diagnósticos da brucelose bovina via Inferência Bayesiana**. 59 p. Dissertação (Mestrado) — Universidade Federal de Viçosa - UFV, Viçosa - Minas Gerais, 2013.
- SPEYBROECK, N. et al. Misclassification errors in prevalence estimation: Bayesian handling with care. **International Journal of Public Health**, v. 58, n. 5, p. 791 – 795, 2013.
- SPICKLER, A. R. **Bovine Brucellosis: Brucella abortus**. 2009a. Disponível em: <<http://www.cfsph.iastate.edu/DiseaseInfo/factsheets.php>>. Acesso em: 10 Set 2015.
- SPICKLER, A. R. **Brucellosis**. 2009b. Disponível em: <<http://www.cfsph.iastate.edu/DiseaseInfo/factsheets.php>>. Acesso em: 01 Set 2015.
- SPIEGELHALTER, D. J. et al. Bayesian measures of model complexity and fit. **Journal of the Royal Statistical Society: Series B (Statistical Methodology)**, v. 64, n. 4, p. 583 – 639, 2002.
- STONE, C. A.; ZHU, X. **Bayesian Analysis of Item Response Theory Models Using SAS**. Cary, North Carolina: SAS Institute, 2015. 280 p.
- STURTZ, S.; LIGGES, U.; GELMAN, A. R2winbugs: A package for running winbugs from r. **Journal of Statistical Software**, v. 12, n. 3, p. 1–16, 2005. Disponível em: <<http://www.jstatsoft.org>>.
- TOFT, N.; JØRGENSEN, E.; HØJSGAARD, S. Diagnosis diagnostic tests: Evaluating the assumptions underlying the estimation of sensitivity and specificity in the absence of a gold standard. **Preventive Veterinary Medicine**, v. 68, p. 19 – 33, 2005.
- TORRANCE-RYNARD, V. L.; WALTER, S. D. Effects of dependence errors in the assessment of diagnostic test performance. **Statistics in Medicine**, v. 16, p. 2157 – 2175, 1997.
- VACEK, P. M. The effect of conditional dependence on the evaluation of diagnostic tests. **Biometrics**, v. 41, n. 4, p. 959 – 968, 1985.
- VERMUNT, J. K.; MAGIDSON, J. Latent variable. In: LEWIS-BECK, M.; BRYMAN, T. L. A. (Ed.). **The Sage encyclopedia of social sciences research methods**. Thousand Oakes: Sage, 2004. p. 555 – 556.
- VIDAL, E. et al. Estimation of the accuracy of two diagnostic methods for the detection of plum pox virus in nursery blocks by latent class models. **Plant Pathology**, v. 61, p. 413 – 422, 2012.

WALTER, S. D.; IRWING, L. M. Estimation of test error rates, disease prevalence and relative risk from misclassified data: A review. **Journal of Clinical Epidemiology**, v. 41, n. 9, p. 923 – 937, 1988.

WANG, Z.; ZHOU, X. **Nonparametric Identifiability of Finite Mixture Models with Covariates for Estimating Error Rate without a Gold Standard**. 2014. 1 - 27 p. Disponível em: <<http://biostats.bepress.com/uwbiostat/paper403>>. Acesso em: 25 out. 2016.

WIERINGEN, W. N. van. On identifiability of certain latent class models. **Statistics & Probability Letters**, v. 75, p. 211 – 218, 2005.

XIE, Y.; CARLIN, B. P. **A Note on Bayesian Learning and Identifiability in Hierarchical Models**. 2004. Disponível em: <<http://www.sph.umn.edu/faculty1/wp-content/uploads/2012/11/rr2004-009.pdf>>. Acesso em: 31 out. 2016.

XU, G.; ZHANG, S. Identifiability of diagnostic classification models. **Psychometrika**, v. 81, n. 3, p. 625 – 649, september 2016.

YANG, I.; BECKER, M. P. Latent variable modeling of diagnostic accuracy. **Biometrics**, v. 53, p. 948 – 958, 1997.

YOUNG, M. A. Evaluating diagnostic criteria: A latent class paradigm. **Journal of Psychiatric Research**, v. 17, n. 3, p. 285 – 296, 1982.

APÊNDICE

A

A seguir apresenta-se: os valores dos critérios de Raftery e Lewis (F_D) e Gelman e Rubin (\hat{R}) para os parâmetros de covariâncias condicionais dos Modelos 2, 3, 4, 5, Tabela 11, e os gráficos dos traços, correlações e densidades dos parâmetros: prevalência (δ), sensibilidades (S) e especificidades (E) para os Modelos 1 e 2. Para os Modelos 3, 4 e 5, resultados semelhantes foram encontrados.

Tabela 11: Critérios de convergência de Raftery e Lewis (F_D) e Gelman e Rubin (\hat{R}).

Parâmetros	Modelo 2		Modelo 3		Modelo 4		Modelo 5	
	F_D	\hat{R}	F_D	\hat{R}	F_D	\hat{R}	F_D	\hat{R}
$cov_{S_{12}}$	0,991	1,000					1,010	1,000
$cov_{E_{12}}$	0,988	1,000					1,010	1,000
$cov_{S_{13}}$			0,986	1,000			0,983	1,000
$cov_{E_{13}}$			0,998	1,000			1,000	1,000
$cov_{S_{23}}$					1,010	1,000	1,010	1,000
$cov_{E_{23}}$					1,020	1,000	0,991	1,000

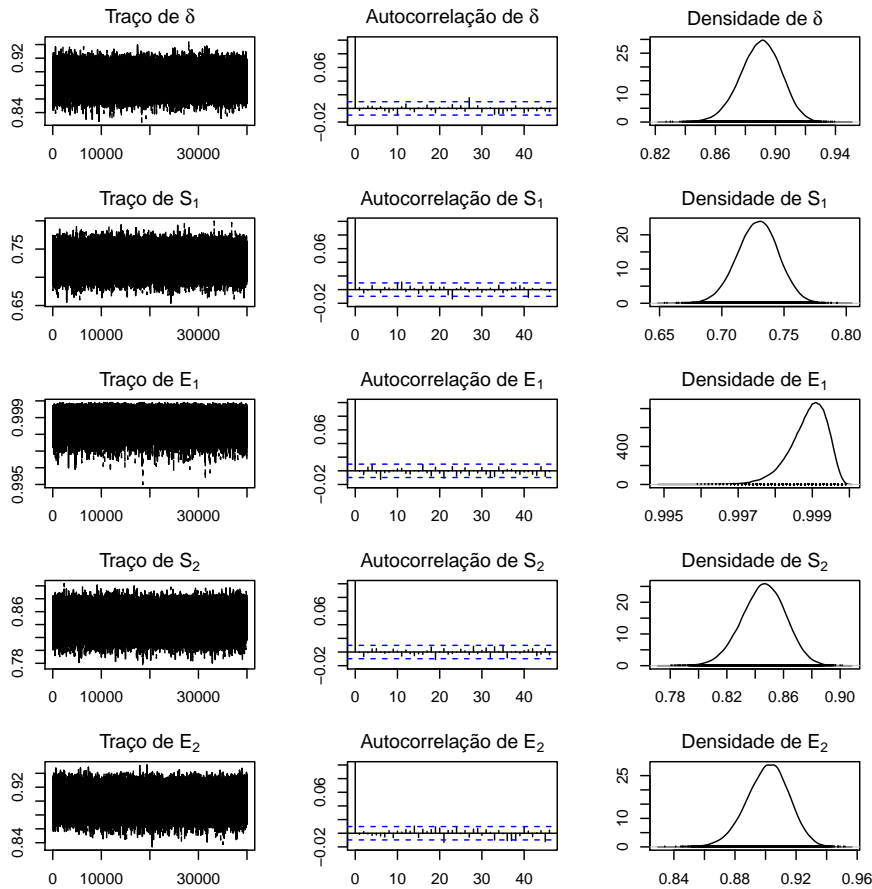


Figura 8: Gráficos dos traços, correlações e densidades dos parâmetros: prevalência (δ), sensibilidades (S) e especificidades (E) para o Modelo 1.

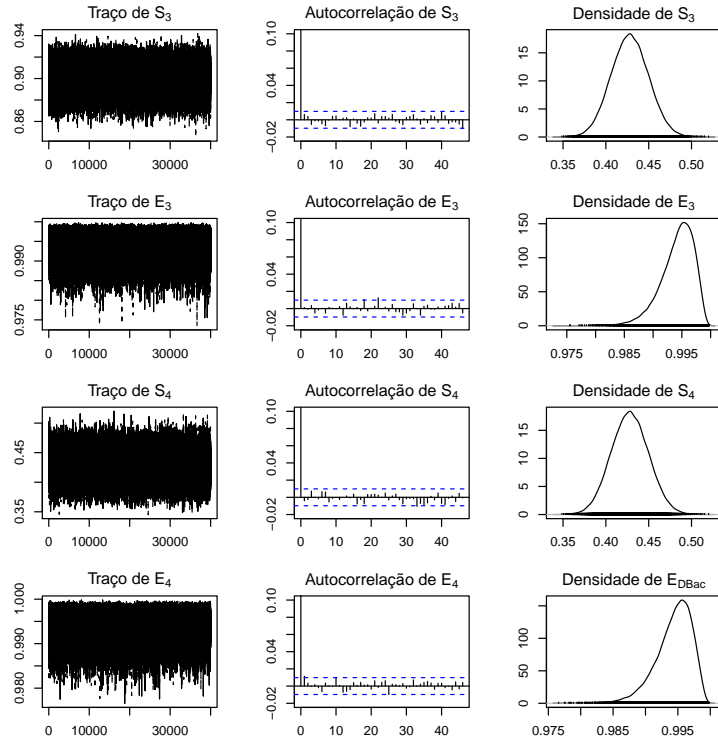


Figura 9: Continuação dos gráficos dos traços, correlações e densidades dos parâmetros: sensibilidade (S) e especificidades (E) para o Modelo 1.

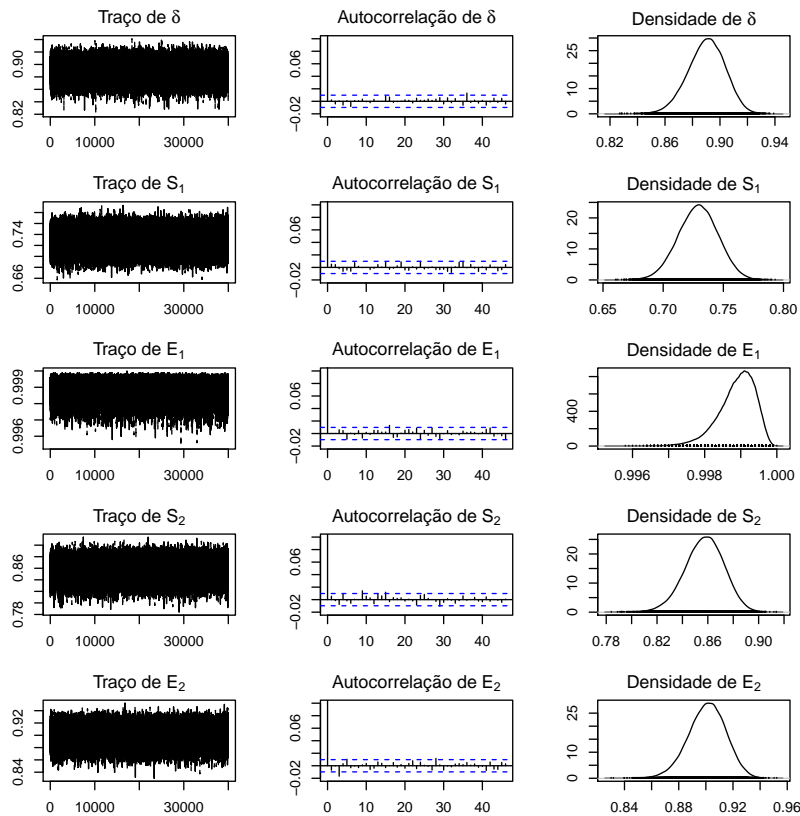


Figura 10: Gráficos dos traços, correlações e densidades dos parâmetros: prevalência (δ), sensibilidade (S) e especificidades (E) para o Modelo 2.

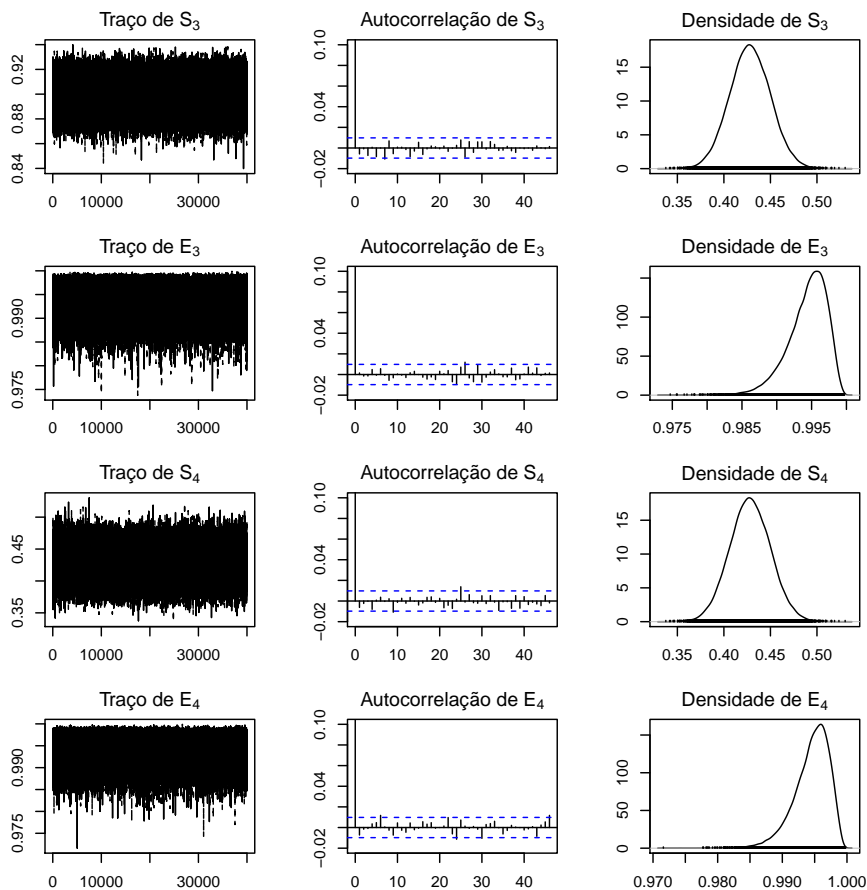


Figura 11: Continuação dos gráficos dos traços, correlações e densidades dos parâmetros: sensibilidades (S) e especificidades (E) para o Modelo 2.

B

Na Tabela 12 estão apresentados os desvios padrão das médias das distribuições a posteriori dos parâmetros: prevalência, sensibilidades, especificidades e correlações dos Testes AAT, 2-ME, FC e DBac, conforme os Modelos 1, 2, 3, 4 e 5.

Tabela 12: Desvios padrão das médias das distribuições a posteriori da prevalência, sensibilidades, especificidades, covariâncias e correlações condicionais dos modelos 1, 2, 3, 4 e 5.

Parâmetros	Desvios Padrão				
	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
δ	0,013	0,013	0,014	0,013	0,014
S_1	0,016	0,016	0,017	0,016	0,017
E_1	0,001	0,001	0,001	0,001	0,001
S_2	0,015	0,015	0,015	0,016	0,016
E_2	0,014	0,014	0,014	0,014	0,014
S_3	0,011	0,011	0,011	0,011	0,012
E_3	0,003	0,003	0,003	0,003	0,003
S_4	0,022	0,022	0,022	0,022	0,022
E_4	0,003	0,003	0,003	0,003	0,004
cov_{S12}		0,007			0,002
cov_{E12}		0,0004			0,0004
cov_{S13}			0,005		0,002
cov_{E13}			0,0004		0,0004
cov_{S23}				0,004	0,002
cov_{E23}				0,002	0,002
ρ_{D12}		0,045			0,015
ρ_{ND12}		0,034			0,033
ρ_{D13}			0,039		0,016
ρ_{ND13}			0,163		0,165
ρ_{D23}				0,039	0,018
ρ_{ND23}				0,073	0,074

C

Códigos do programa R para os Modelos 1 e 2.

```
# Testes diagnósticos: AAT, 2-ME, FC e DBac
# Considerando a suposição de independência condicional

sink("modelo1.txt")
cat(" model {

# Função de verossimilhança

Vero[1:16] ~ dmulti(p[1:16],N)

p[1 ] <- P*( (S[1])* (S[2])* (S[3])* (S[4]) ) + (1-P)*( (1-E[1])* (1-E[2])* (1-E[3])* (1-E[4]) )
p[2 ] <- P*( (S[1])* (S[2])* (S[3])* (1-S[4]) ) + (1-P)*( (1-E[1])* (1-E[2])* (1-E[3])* (E[4]) )
p[3 ] <- P*( (S[1])* (S[2])* (1-S[3])* (S[4]) ) + (1-P)*( (1-E[1])* (1-E[2])* (E[3])* (1-E[4]) )
p[4 ] <- P*( (S[1])* (S[2])* (1-S[3])* (1-S[4]) ) + (1-P)*( (1-E[1])* (1-E[2])* (E[3])* (E[4]) )
p[5 ] <- P*( (S[1])* (1-S[2])* (S[3])* (S[4]) ) + (1-P)*( (1-E[1])* (E[2])* (1-E[3])* (1-E[4]) )
p[6 ] <- P*( (S[1])* (1-S[2])* (S[3])* (1-S[4]) ) + (1-P)*( (1-E[1])* (E[2])* (1-E[3])* (E[4]) )
p[7 ] <- P*( (S[1])* (1-S[2])* (1-S[3])* (S[4]) ) + (1-P)*( (1-E[1])* (E[2])* (E[3])* (1-E[4]) )
p[8 ] <- P*( (S[1])* (1-S[2])* (1-S[3])* (1-S[4]) ) + (1-P)*( (1-E[1])* (E[2])* (E[3])* (E[4]) )
```

```

p[9 ] <- P*( (1-S[1])*S[2])*S[3]*S[4] ) + (1-P)*( (E[1])*(1-E[2])*(1-E[3])*(1-E[4]) )
p[10] <- P*( (1-S[1])*S[2])*S[3]*(1-S[4]) ) + (1-P)*( (E[1])*(1-E[2])*(1-E[3])*(E[4]) )
p[11] <- P*( (1-S[1])*S[2])*(1-S[3])*S[4] ) + (1-P)*( (E[1])*(1-E[2])*(E[3])*(1-E[4]) )
p[12] <- P*( (1-S[1])*S[2])*(1-S[3])*(1-S[4]) ) + (1-P)*( (E[1])*(1-E[2])*(E[3])*(E[4]) )
p[13] <- P*( (1-S[1])*(1-S[2])*S[3])*S[4] ) + (1-P)*( (E[1])*(E[2])*(1-E[3])*(1-E[4]) )
p[14] <- P*( (1-S[1])*(1-S[2])*S[3])*(1-S[4]) ) + (1-P)*( (E[1])*(E[2])*(1-E[3])*(E[4]) )
p[15] <- P*( (1-S[1])*(1-S[2])*(1-S[3])*S[4] ) + (1-P)*( (E[1])*(E[2])*(E[3])*(1-E[4]) )
p[16] <- P*( (1-S[1])*(1-S[2])*(1-S[3])*(1-S[4]) ) + (1-P)*( (E[1])*(E[2])*(E[3])*(E[4]) )

```

Distribuições a priori das sensibilidades (S), especificidades (E) e prevalência (P)

```

P ~ dbeta(119.3, 51.13)
S[1] ~ dbeta(315.4506,7.756983)
E[1] ~ dbeta(4389.326,4.833576)
S[2] ~ dbeta(118.22, 15.36)
E[2] ~ dbeta(454.5354, 42.2246)
S[3] ~ dbeta(297.0639,15.47045)
E[3] ~ dbeta(791.025,3.975)
S[4] ~ dbeta(49.5, 49.5)
E[4] ~ dbeta(791.03, 3.98)

```

```

}"; fill = T)

```

```

sink()

```

```

# Testes diagnósticos: AAT, 2-ME, FC e DBac
# Considerando dependencia condicional entre AAT e 2-ME

```

```

sink("modelo2.txt")
cat(" model {

```

```

# Função de verossimilhança

```

```

Vero[1:16] ~ dmulti(p[1:16],N)

```

```

p[1 ] <- P*((S[1])*S[2]) + covs12)*((S[3])*S[4]) + (1-P)*((1-E[1])*(1-E[2]) + cove12)*( (1-E[3])*(1-E[4]) )
p[2 ] <- P*((S[1])*S[2]) + covs12)*((S[3])*(1-S[4])) + (1-P)*((1-E[1])*(1-E[2]) + cove12)*( (1-E[3])*(E[4]) )
p[3 ] <- P*((S[1])*S[2]) + covs12)*((1-S[3])*S[4]) + (1-P)*((1-E[1])*(1-E[2]) + cove12)*( (E[3])*(1-E[4]) )
p[4 ] <- P*((S[1])*S[2]) + covs12)*((1-S[3])*(1-S[4])) + (1-P)*((1-E[1])*(1-E[2]) + cove12)*( (E[3])*(E[4]) )
p[5 ] <- P*((S[1])*(1-S[2]) - covs12)*((S[3])*S[4]) + (1-P)*((1-E[1])*(E[2]) - cove12)*( (1-E[3])*(1-E[4]) )
p[6 ] <- P*((S[1])*(1-S[2]) - covs12)*((S[3])*(1-S[4])) + (1-P)*((1-E[1])*(E[2]) - cove12)*( (1-E[3])*(E[4]) )
p[7 ] <- P*((S[1])*(1-S[2]) - covs12)*((1-S[3])*S[4]) + (1-P)*((1-E[1])*(E[2]) - cove12)*( (E[3])*(1-E[4]) )
p[8 ] <- P*((S[1])*(1-S[2]) - covs12)*((1-S[3])*(1-S[4])) + (1-P)*((1-E[1])*(E[2]) - cove12)*( (E[3])*(E[4]) )
p[9 ] <- P*((1-S[1])*S[2]) - covs12)*((S[3])*S[4]) + (1-P)*((E[1])*(1-E[2]) - cove12)*( (1-E[3])*(1-E[4]) )
p[10] <- P*((1-S[1])*S[2]) - covs12)*((S[3])*(1-S[4])) + (1-P)*((E[1])*(1-E[2]) - cove12)*( (1-E[3])*(E[4]) )
p[11] <- P*((1-S[1])*S[2]) - covs12)*((1-S[3])*S[4]) + (1-P)*((E[1])*(1-E[2]) - cove12)*( (E[3])*(1-E[4]) )
p[12] <- P*((1-S[1])*S[2]) - covs12)*((1-S[3])*(1-S[4])) + (1-P)*((E[1])*(1-E[2]) - cove12)*( (E[3])*(E[4]) )
p[13] <- P*((1-S[1])*(1-S[2]) + covs12)*((S[3])*S[4]) + (1-P)*((E[1])*(E[2]) + cove12)*( (1-E[3])*(1-E[4]) )
p[14] <- P*((1-S[1])*(1-S[2]) + covs12)*((S[3])*(1-S[4])) + (1-P)*((E[1])*(E[2]) + cove12)*( (1-E[3])*(E[4]) )
p[15] <- P*((1-S[1])*(1-S[2]) + covs12)*((1-S[3])*S[4]) + (1-P)*((E[1])*(E[2]) + cove12)*( (E[3])*(1-E[4]) )

```

```

p[16] <- P*((1-S[1])*(1-S[2]) + covs12)*((1-S[3])*(1-S[4])) + (1-P)*((E[1])*(E[2]) + cove12)*((E[3])*(E[4]))

# Distribuições a priori das sensibilidades (S), especificidades (E) e prevalência (P)

P ~ dbeta(119.3, 51.13)
S[1] ~ dbeta(315.4506,7.756983)
E[1] ~ dbeta(4389.326,4.833576)
S[2] ~ dbeta(118.22, 15.36)
E[2] ~ dbeta(454.5354, 42.2246)
S[3] ~ dbeta(297.0639,15.47045)
E[3] ~ dbeta(791.025,3.975)
S[4] ~ dbeta(49.5, 49.5)
E[4] ~ dbeta(791.03, 3.98)

# Dependência condicional entre os testes AAT x 2-ME

# Intervalos para as covariâncias condicionais

ls12 <- (S[1] - 1)*(1 - S[2])
Ls12 <- min(S[1],S[2]) - S[1]*S[2]

le12 <- (E[1] - 1)*(1 - E[2])
Le12 <- min(E[1],E[2]) - E[1]*E[2]

# Covariâncias condicionais

covs12 ~ dunif(ls12,Ls12)
cove12 ~ dunif(le12,Le12)

# Correlações condicionais
rhos12 <- covs12/sqrt( S[1]*(1-S[1])*S[2]*(1-S[2]) )
rhoe12 <- cove12/sqrt( E[1]*(1-E[1])*E[2]*(1-E[2]) )

)", fill = T)
sink()

```