PARECER

Ribeirão Preto, 11/01/2017.

A aluna Ananda Amorim Vieira desenvolveu seu trabalho sob minha orientação com o tema **“Desenvolvimento e Implantação de Sistema para a Gestão da Informação e Vigilância de Anomalias Congênitas no Brasil”**. Neste projeto tem-se por objetivo principal apresentar o sistema de vigilância de anomalias congênitas (AC) utilizando três bases públicas do Sistema Único de Saúde (Sistema de Informações sobre Nascidos Vivos - SINASC, Sistema de Informações Hospitalares - SIH e Sistema de Informações sobre Mortalidade - SIM) com intuito de monitorar as anomalias no Brasil mapeando os nascimentos, internações e óbitos. Este sistema tornou-se o Portal de Monitoramento de AC (PMAC), com dados que abrangem o período de janeiro 2001 a dezembro 2011, conta com ferramentas que correspondem aos módulos de auditoria automática que verifica subnotificações e inconsistência, módulos de análises epidemiológicas contendo incidência, fluxo de migração e coeficientes de mortalidade (infantil, perinatal, neonatal e segundo causas por AC). Este estudo também trás a realização da técnica do relacionamento probabilístico com a integração das três bases contribuindo com informações de rastreabilidade dos registros com anomalias.

A aluna teve participação ativa, com excelente desempenho, em todas as etapas de desenvolvimento do projeto, da revisão bibliográfica à análise e discussão dos resultados, passando pela escolha da abordagem metodológica e de sua aplicação.

Assim, frente ao exposto, considero a aluna apta para apresentação de sua dissertação ao programa de Pós-Graduação em Saúde na Comunidade do Departamento de Medicina Social da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo, para a aquisição do título de Mestre em Ciências.

Domingos Alves
Departamento de Medicina Social
Faculdade de Medicina de Ribeirão Preto – Universidade de São Paulo

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE MEDICINA DE RIBEIRÃO PRETO
DEPARTAMENTO DE MEDICINA SOCIAL

ANANDA AMORIM VIEIRA

**Desenvolvimento e Implantação de Sistema para a Gestão da
Informação e Vigilância de Anomalias Congênitas no Brasil**

Ribeirão Preto

2017

ANANDA AMORIM VIEIRA

**Desenvolvimento e Implantação de Sistema para a Gestão da Informação e Vigilância
de Anomalias Congênitas no Brasil**

Dissertação apresentada ao Programa de Pós-Graduação em Saúde na Comunidade do Departamento de Medicina Social da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo, como parte das exigências para a obtenção do título de Mestre em Ciências

Orientador: Prof. Dr. Domingos Alves

Ribeirão Preto

2017

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Catálogo da Publicação

Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo

Vieira, Ananda Amorim

Desenvolvimento e Implantação de Sistema para a Gestão da Informação e Vigilância de Anomalias Congênitas no Brasil.

Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo. Ribeirão Preto, São Paulo, 2017.

114 f. : il.

Dissertação de Mestrado, apresentada ao Programa de Pós-graduação em Saúde na Comunidade do Departamento de Medicina Social da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo.

Orientador: Domingos Alves

1. Sistema de Informação em Saúde. 2. Anomalias Congênitas. 3. Portal de Conteúdo em Saúde. 4. Linkage Probabilístico.

Nome: VIEIRA, Ananda Amorim

Título: Desenvolvimento e Implantação de Sistema para a Gestão da Informação e Vigilância de Anomalias Congênitas no Brasil

Dissertação apresentada ao Programa de Pós-Graduação em Saúde na Comunidade do Departamento de Medicina Social da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo, como parte das exigências para a obtenção do título de Mestre em Ciências.

Aprovado em: ____/____/____

Banca Examinadora

Prof.(a) Dr.(a)_____

Instituição _____

Assinatura_____

Prof.(a) Dr.(a)_____

Instituição _____

Assinatura_____

Prof.(a) Dr.(a)_____

Instituição _____

Assinatura_____

DEDICATÓRIA

A minha família, e em especial ao meu querido irmão Lucas Amorim Vieira que conviveu comigo por vinte e um anos e que sempre sentirei saudades.

AGRADECIMENTOS

Primeiramente gostaria de agradecer a todos os meus professores desde o ensino fundamental até os do ensino superior, pois a educação é o primeiro passo para o desenvolvimento pessoal e intelectual.

Ao orientador Dr. Domingos Alves e ao professor Dr. Victor Evangelista de Faria Ferraz, que muito contribuíram com meu crescimento científico, pelo excelente trabalho como pesquisadores e pela confiança em mim concedida na realização deste projeto.

Ao Programa de Pós - Graduação em Saúde na Comunidade do Departamento de Medicina Social da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo, pela oportunidade de realização do mestrado.

As amigas de graduação Nathalia e Sara, que auxiliaram em etapas de teste deste projeto e que são exemplos de pessoas boas que se deve cultivar para o resto da vida.

A amiga de graduação e pós-graduação Lívia, pelos momentos de convivência no trabalho, nas aulas e que é um exemplo de perseverança e admiração como workaholic.

Aos companheiros e profissionais do laboratório grupo ORAH, pelo suporte em manter o site PMAC hospedado no servidor e pelas horas de trocas de experiência e convivência.

Aos amigos do Karatê Dionne, Juliana, Sara, Reginaldo, Solimar e sensei Oswaldo, pois através do esporte adquiri mais tranquilidade para vencer as dificuldades do dia a dia.

As minhas sempre amigas Andréia, Leticia, Liliane, Marcinha e Priscila, pelas risadas e diversos momentos de descontração necessária para manter o equilíbrio mental.

Aos meus familiares, em especial ao meu tio Carlos Alberto, que sempre me incentivou e auxiliou na continuação dos estudos, além de ser exemplo de dedicação.

Aos meus queridos pais Edson e Nelcimar por todo o incentivo, amor, paciência, compreensão e apoio que são dedicadas a mim todos os dias, bem como ao meu namorado Paulo Augusto que emprega seu carinho e atenção constantemente.

E finalmente ao bom Deus que me proporcionou tudo que tenho hoje.

RESUMO

VIEIRA, A. A. **Desenvolvimento e implantação de sistema para a gestão da informação e vigilância de anomalias congênitas no Brasil**. 2017. 114 f. Dissertação (Mestrado) – Programa de Pós-Graduação em Saúde na Comunidade, Departamento de Medicina Social, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2017.

Anomalia congênita (AC) é causada por alterações morfológicas e/ou funcionais ocasionada no desenvolvimento embrionário ou fetal devido a fatores genéticos, ambientais, multifatoriais ou desconhecidos. A detecção pode ser no período pré-natal, no recém-nascido ou posteriormente. Estudos recentes indicam que de 2% a 3% de todos os nascidos vivos no mundo apresentam AC. Alguns autores afirmam que no Brasil elas constituem a segunda causa de mortalidade infantil, com 11,2%. Particularmente, o país conta com sistemas de bancos de dados em saúde com informações sobre anomalias congênitas. Dentre estes sistemas, temos três que se destacam: o Sistema de Informações Hospitalares do SUS (SIH), o Sistema de Informações sobre Nascidos Vivos (SINASC) e o Sistema de Informações sobre Mortalidade (SIM). Todos são obtidos através do site do Departamento de Informática do Sistema Único de Saúde (DATASUS). Todavia, esses sistemas, de maneira isolada, não fornecem uma visão clara tanto da epidemiologia das anomalias congênitas, quanto dos indivíduos com esse agravo, se ao terem nascidos foram tratados ou em decorrência dela faleceram. Diante deste cenário, neste projeto tem-se por objetivo principal apresentar o sistema de vigilância de AC utilizando estas bases públicas com intuito de monitorar as anomalias no Brasil, mapeando os nascimentos, internações e óbitos. Este sistema tornou-se o Portal de Monitoramento de AC (PMAC), com dados que abrangem o período de janeiro 2001 a dezembro 2011. Ele conta com ferramentas que correspondem aos módulos de auditoria automática, que verifica subnotificações e inconsistência, módulos de análises epidemiológicas contendo incidência, fluxo de migração e coeficientes de mortalidade (infantil, perinatal, neonatal e segundo causas por AC). Este estudo também traz a realização da técnica do relacionamento probabilístico com a integração dos sistemas SINASC, SIH e SIM, que formaram um banco contribuindo com informações de rastreabilidade dos registros com anomalias, além de apresentar os resultados satisfatórios dos testes caixa-preta junto com a aplicação da técnica de avaliação heurística, realizados por usuários para avaliar a usabilidade do portal. Por fim, ressalta-se que o portal é dinâmico, e a manipulação dos dados por meio do acesso dos usuários a análises de cunho epidemiológico, em diversas agregações, permite auxiliar a gestão das informações, promover reflexões em torno desta temática, incentivar mais estudos e pesquisas, bem como prover recursos e apoiar medidas sociais e políticas para assistir os indivíduos acometidos por AC, suas famílias e a comunidade.

Palavras-chave: Sistema de informação em saúde; Anomalias congênitas; Portal de conteúdo em saúde; Linkage probabilístico.

ABSTRACT

VIEIRA, A. A. **Development and implementation of a system for information management and surveillance of congenital anomalies in Brazil**. 2017. 114 f. Dissertação (Mestrado) – Programa de Pós-Graduação em Saúde na Comunidade, Departamento de Medicina Social, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2017.

Congenital anomaly (AC) is caused by morphological and / or functional alterations originated during embryonic or fetal development due to genetic, environmental, multifactorial or unknown factors. Detection may happen in the prenatal period, in the newborn or later. Recent studies indicate that AC is present in about 2% to 3% of all live births worldwide. Some authors affirm that in Brazil they are the second cause of infant mortality, with around 11.2%. In particular, Brazil has health databases systems with information about congenital anomalies. Among these systems we have three that stand out: the Hospital Information System of SUS (SIH), the Live Birth Information System (SINASC) and Mortality Information System (SIM). All of which can be obtained through the Informatics of public Health System Department's website (DATASUS). However, these systems in isolation do not provide a clear view both of the epidemiology of congenital anomalies as well as the individuals with this birth defect be treated or it result of their death. In this scenario, the main objective of this project is to present an AC surveillance system that uses these public databases to monitor the anomalies in Brazil by mapping births, hospitalizations and deaths. This system became the Portal of Monitoring of AC (PMAC), with data that cover the period from January 2001 to December 2011, with tools that correspond to the modules of automatic audit that verifies under notifications and inconsistency, modules of epidemiological analyzes containing incidence, migration flow and mortality coefficients (infant, perinatal, neonatal and second AC causes). This study also used the probabilistic linkage technique with an integration of the SINASC, SIH and SIM systems to build a database contributing with traceable information for the records with anomalies. In addition is presented satisfactory results for black box tests along with the portal usability evaluation performed with users through heuristic evaluation technique. Finally, it should be noted that the portal is dynamic offering options like user access to data manipulation through epidemiological analyzes in diverse aggregations. It helps to manage information, promote reflections around this theme, encourage further studies and research, as well as provide resources and support social and political measures to assist individuals affected by AC, their families and the community.

Keywords: Health information system; Congenital anomalies; Health content portal; Probabilistic Linkage.

LISTA DE FIGURAS

Figura 1 – Representação dos menus e submenus presentes no portal PMAC	10
Figura 2 – Representação da classificação segundo os graus de avaliação que vai de 0% a 100%	17
Figura 3 – Representação ilustrativa da Padronização e Relacionamento	22
Figura 4 – Representação da evolução da quantidade de campos nas bases de janeiro de 2001 a dezembro de 2011	32
Figura 5 – Representação da evolução da quantidade de registros inseridos no SINASC de janeiro de 2001 a dezembro de 2011.	33
Figura 6 – Representação da evolução da quantidade de registros inseridos no SIH de janeiro de 2001 a dezembro de 2011.	34
Figura 7 – Representação da evolução da quantidade de registros inseridos no SIM de janeiro de 2001 a dezembro de 2011.	35
Figura 8 – Representação da incompletude da variável id anomalia no SINASC de janeiro de 2001 a dezembro de 2011.....	39
Figura 9 – Inconsistência nos campos com notificação de AC na base SINASC de janeiro de 2001 a dezembro de 2011.....	44
Figura 10 – Coeficiente de Incidência de notificação de AC na base do SINASC de janeiro de 2001 a dezembro de 2011.....	46
Figura 11 – Coeficiente de Mortalidade Infantil na base do SIM de janeiro de 2001 a dezembro de 2011.....	47
Figura 12 – Taxa de Mortalidade Perinatal na base do SIM de janeiro de 2001 a dezembro de 2011.	48
Figura 13 – Taxa de Mortalidade Neonatal Tardia na base do SIM de janeiro de 2001 a dezembro de 2011.....	49

Figura 14 – Mortalidade Segundo Causas de AC na base do SIM de janeiro de 2001 a dezembro de 2011.....	50
Figura 15 – Exemplo de arquivo gerado pelo software openRecLink com os cálculos da etapa do relacionamento.....	52
Figura 16 – Ilustração da etapa combinada gerado pelo software openRecLink.	53
Figura 17 – Relacionamento probabilístico SIHAC-SINASCAC.....	54
Figura 18 – Relacionamento probabilístico SIMAC-SINASCAC.....	55
Figura 19 – Interface do menu Início do PMAC	58
Figura 20 – Interface do submenu Histórico de Informações Gerais.....	59
Figura 21 – Interface do submenu Histórico SIHAC.....	59
Figura 22 – Interface do submenu Histórico SINASC.....	60
Figura 23 – Interface do submenu Histórico SIM.....	60
Figura 24 – Interface do submenu Histórico do Linkage.....	60
Figura 25 – Interface do submenu Busca de Dados no banco SIHAC	61
Figura 26 – Interface dos resultados do submenu Busca de Dados no banco SIHAC	62
Figura 27 – Interface do submenu Auditoria Automática no banco SINASC	62
Figura 28 – Interface dos resultados do submenu Auditoria Automática no banco SINASC.	63
Figura 29 – Interface do submenu Frequência no banco SIM.....	63
Figura 30 – Interface dos resultados do submenu Frequência no banco SIM	64
Figura 31 – Interface do submenu Fluxo dos Nascidos Vivos no banco SINASC	65
Figura 32 – Interface dos resultados do submenu Fluxo dos Nascidos Vivos no banco SINASC.....	66
Figura 33 – Interface do submenu Incidência no banco SINASC.....	67

Figura 34 – Interface dos resultados do submenu Incidência no banco SINASC	67
Figura 35 – Interface do submenu Inconsistência no banco SINASC.....	68
Figura 36 – Interface dos resultados do submenu Inconsistência no banco SINASC	68
Figura 37 – Interface do submenu Fluxo de Mortalidade no banco SIM	69
Figura 38 – Interface dos resultados do submenu Fluxo de Mortalidade no banco SIM.....	70
Figura 39 – Interface do submenu Coeficiente de Mortalidade no banco SIM.....	71
Figura 40 – Interface dos resultados do submenu Coeficiente de Mortalidade no banco SIM.....	71
Figura 41 – Interface do submenu Mortalidade Segundo Causas no banco SIM.....	72
Figura 42 – Interface dos resultados do submenu Mortalidade Segundo Causas no banco SIM.	72
Figura 43 – Interface do menu Links.....	73
Figura 44 – Interface do submenu Equipe.....	74
Figura 45 – Interface do submenu Contato	74
Figura 46 – Interface do submenu Pesquisa de Satisfação	75

LISTA DE TABELAS

Tabela 1 – Padronização da variável “sexo”	24
Tabela 2 – Padronização das variáveis “data”	24
Tabela 3 – Padronização da variável “código município”	24
Tabela 4 – Padronização das variáveis com valor “CID-10 (cap. XVII)”	25
Tabela 5 – Padronização das variáveis “código do estado”	25
Tabela 6 – Multiplicação entre arquivos relacionados SINASC-SIHAC	27
Tabela 7 – Multiplicação entre arquivos relacionados SINASC-SIM	28
Tabela 8 – Chaves de blocagem	29
Tabela 9 – Variáveis utilizadas no pareamento SINASC - SIM com o algoritmo de comparação	30
Tabela 10 – Variáveis utilizadas no pareamento SINASC - SIHAC com o algoritmo de comparação	31
Tabela 11 – Porcentagem da incompletude referente a todos os dados disponibilizados do SINASC no período de janeiro 2001 a dezembro 2011	37
Tabela 12 – Continuação da tabela 1 referente a porcentagem da incompletude referente a todos os dados disponibilizados do SINASC no período de janeiro 2001 a dezembro	38
Tabela 13 – Porcentagem da incompletude referente a todos os dados disponibilizados do SIM no período de janeiro 2001 a dezembro 2011	40
Tabela 14 – Continuação da tabela 3 referente a porcentagem da incompletude referente a todos os dados disponibilizados do SIM no período de janeiro 2001 a dezembro	41
Tabela 15 – Porcentagem da incompletude referente as variáveis disponibilizadas pelo SIHAC no período de janeiro 2001 a dezembro 2011	43
Tabela 16 – Resultados da rastreabilidade dos dados do SIHAC, SINASCAC e SIMAC	56

Tabela 17 – Continuação dos resultados da rastreabilidade dos dados dos linkage SIHAC-SINASCAC e SIMAC-SINASCAC	57
Tabela 18 – Avaliação heurística dos especialistas	76
Tabela 19 – Avaliação heurística dos usuários	77

LISTA DE EQUAÇÕES

Equação 1 – Fórmula de incompletude das variáveis.....	16
Equação 2 – Fórmula de inconsistência do campo de identificação de AC no SINASC.....	17
Equação 3 – Fórmula do coeficiente de incidência	18
Equação 4 – Fórmula do coeficiente de incidência para AC	19
Equação 5 – Fórmula do coeficiente de mortalidade infantil	19
Equação 6 – Fórmula da taxa de mortalidade perinatal	20
Equação 7 – Fórmula da taxa de mortalidade neonatal tardia	20
Equação 8 – Fórmula da mortalidade segundo causas	21
Equação 9 – Fórmula de concordância das variáveis	29
Equação 10 – Fórmula da discordância das variáveis	30

LISTA DE ABREVIATURAS E SIGLAS

AC	Anomalias Congênitas
AIH	Autorização de Internação Hospitalar
CSS	Cascading Style Sheets
CENEPI	Centro Nacional de Epidemiologia
CID10	Classificação Internacional de Doenças
DN	Declaração de Nascido Vivo
DO	Declaração de Óbito
DATASUS	Departamento de Informática do Sistema Único de Saúde
DENASUS	Departamento Nacional de Auditoria do SUS
ECLAMC	Estudo Colaborativo Latino – Americano de Malformações Congênitas
FUNASA	Fundação Nacional de Saúde
HTML	HyperText Markup Language
INAMPS	Instituto Nacional de Assistência Médica e Previdência Social
PHP	Hypertext Preprocessor
PMAC	Portal de Monitoramento de Anomalia Congênita
SGBD	Sistema Gerenciador de Banco de Dados
SIH	Sistema de Informações Hospitalares do Sistema Único de Saúde
SIM	Sistema de Informações sobre Mortalidade
SINASC	Sistema de Informações sobre Nascidos Vivos
SNA	Sistema Nacional de Auditoria
SUS	Sistema Único de Saúde

SUMÁRIO

1.	<u>INTRODUÇÃO</u>	<i>1</i>
1.1.	ANOMALIA CONGÊNITA	<i>1</i>
1.2.	SISTEMAS DE INFORMAÇÃO NO BRASIL: SINASC, SIH E SIM	<i>3</i>
1.3.	MOTIVAÇÃO E JUSTIFICATIVA	<i>6</i>
2.	<u>OBJETIVO</u>	<i>7</i>
2.1.	GERAL	<i>7</i>
2.2.	SECUNDÁRIO	<i>7</i>
3.	<u>MATERIAIS E MÉTODOS</u>	<i>8</i>
3.1.	TIPO DE ESTUDO	<i>8</i>
3.2.	FONTE DE DADOS	<i>8</i>
3.3.	BANCO DE DADOS	<i>8</i>
3.4.	DESENVOLVIMENTO DO PORTAL	<i>9</i>
3.4.1.	REQUISITOS DO PORTAL	<i>11</i>
3.4.2.	RECURSOS DE DESENVOLVIMENTO DAS INTERFACES GRÁFICAS.....	<i>12</i>
3.4.3.	ETAPA DE TESTE E AVALIAÇÃO	<i>13</i>
3.5.	AUDITORIA	<i>15</i>
3.5.1.	EVOLUÇÃO DOS BANCOS E DOS DADOS.....	<i>16</i>
3.5.2.	INCOMPLETUDE.....	<i>16</i>
3.5.3.	INCONSISTÊNCIA	<i>17</i>
3.6.	COEFICIENTES EM EPIDEMIOLOGIA	<i>18</i>
3.6.1.	COEFICIENTE DE INCIDÊNCIA	<i>18</i>
3.6.2.	COEFICIENTE DE MORTALIDADE INFANTIL.....	<i>19</i>
3.6.3.	TAXA DE MORTALIDADE PERINATAL	<i>19</i>
3.6.4.	TAXA DE MORTALIDADE NEONATAL TARDIA	<i>20</i>
3.6.5.	MORTALIDADE SEGUNDO CAUSAS.....	<i>20</i>
3.7.	RELACIONAMENTO PROBABILÍSTICO	<i>21</i>
3.7.1.	PADRONIZAÇÃO.....	<i>23</i>

3.7.1.1.	VARIÁVEIS COM CAMPOS EM COMUM ENTRE OS SISTEMAS	23
3.7.1.2.	VARIÁVEIS COM CAMPOS DIFERENTES ENTRE OS SISTEMAS	25
3.7.2.	RELACIONAMENTO	26
3.7.2.1.	BLOCAGEM	26
3.7.2.2.	PAREAMENTO	29
3.8.	ASPECTOS ÉTICOS	31
4.	<u>RESULTADOS</u>	32
4.1.	AUDITORIA	32
4.1.1.	EVOLUÇÃO DOS BANCOS E DOS DADOS	32
4.1.2.	INCOMPLETUDE	35
4.1.2.1.	SINASC	36
4.1.2.2.	SIM	39
4.1.2.3.	SIHAC	42
4.1.3.	INCONSISTÊNCIA	43
4.2.	COEFICIENTES EM EPIDEMIOLOGIA	45
4.2.1.	COEFICIENTE DE INCIDÊNCIA	45
4.2.2.	COEFICIENTE DE MORTALIDADE INFANTIL	46
4.2.3.	TAXA DE MORTALIDADE PERINATAL	47
4.2.4.	TAXA DE MORTALIDADE NEONATAL TARDIA	48
4.2.5.	MORTALIDADE SEGUNDO CAUSAS	49
4.3.	RELACIONAMENTO PROBABILÍSTICO	50
4.4.	PORTAL DE MONITORAMENTO DE ANOMALIA CONGÊNITA	58
4.4.1.	APRESENTAÇÃO DO PMAC	58
4.4.2.	MÓDULO HISTÓRICO	59
4.4.3.	FERRAMENTAS SEMELHANTES ENTRE SINASC, SIHAC E SIM	60
4.4.4.	FERRAMENTAS DISTINTAS NO SINASC	64
4.4.5.	FERRAMENTAS DISTINTAS NO SIM	68
4.4.6.	MÓDULO LINKS	73
4.4.7.	MÓDULO QUEM SOMOS	73
4.4.8.	TESTE DE AVALIAÇÃO	76
5.	<u>DISCUSSÃO</u>	78

6. <u>CONCLUSÃO</u>	81
REFERÊNCIAS	82
ANEXOS	86

1. INTRODUÇÃO

1.1. ANOMALIA CONGÊNITA

As anomalias congênitas (AC), também conhecidas por defeitos congênitos, têm como causa alterações morfológicas e/ou funcionais ocasionadas no desenvolvimento embrionário ou fetal em decorrência de fatores genéticos, ambientais, multifatoriais (junção de ambiental e genética) ou desconhecidos (OLIVEIRA et al., 2007). A detecção pode ser durante o pré-natal, estar aparente ao nascimento ou ficar evidente posteriormente.

As ACs são classificadas quanto à fisiopatogenia por: *malformativos*: defeito morfológico de um órgão ou parte dele gerado por uma anormalidade no processo do desenvolvimento; *deformativos*: através de forças mecânicas externas no desenvolvimento afetando a forma ou posição de parte do corpo do embrião ou feto; *disruptivos*: através de forças mecânicas externas ocasionando defeito morfológico de um órgão, parte dele ou uma região maior do corpo; *displásicos*: organização anormal das células dos tecidos e seus resultados morfológicos. Podendo ser isoladas (única anomalia) ou associativas (mais de uma anomalia) (SÃO PAULO, 2012).

Horovitz, Llerena e Mattos (2005), em seu estudo, indica que no Brasil, por volta dos anos 2000, as ACs chegaram à segunda colocação como causa de óbito e sugere que em parte este dado se deve pela falta de políticas governamentais que atuem na prevenção e manejo deste grupo de indivíduos. A autora evidência algumas medidas governamentais e não governamentais como iniciativas de auxílio e, por estas não serem integradas umas com as outras, acabam por não trazer resultados favoráveis à atenção e acompanhamento das ACs no país. Particularmente, Victora e Barros (2001) apontam que as ACs constituem a segunda causa de mortalidade infantil, com 11,2%.

Estudos recentes indicam que de 2% a 3% de todos os nascidos vivos no mundo apresentam AC. Destes, aproximadamente 60% são de causas desconhecidas. As que são frequentemente conhecidas são as de etiologia genéticas, sendo 6% cromossômicas e 20% de malformações com herança mendeliana. Aquelas relacionadas a fatores ambientais são mais difíceis de detecção e mais numerosas (NICOLA et al., 2010).

Dentre os fatores de risco que podem aumentar a probabilidade do embrião ou feto ser acometido por uma anomalia e que mais se destacam são: a gestante ingerir álcool, fármacos

ou drogas, ter contato com radiação, uma alimentação ruim ou precária, contrair uma doença ou infecção (como a rubéola ou a sífilis congênita), sofrer traumas físicos, possuir algum transtorno hereditário, ter condições socioeconômicas desfavoráveis e estar em uma faixa etária de risco, com relação à idade da mãe – estudos na Inglaterra mostram uma maior frequência de recém-nascido com anomalia congênita em mulheres que possuam mais de 35 anos (BRITO et al., 2010).

No estudo de Schuller-Faccini et al. (2002) há citações de fatores de riscos, dos quais países em desenvolvimento contêm características que podem acentuar os problemas que influenciam na gestação comparadas com países desenvolvidos, como por exemplo: alta incidência de doenças infecciosas, baixa atenção a saúde e pesquisas, práticas frequentes de automedicação e proibição legal da interrupção da gestação.

O acompanhamento do desenvolvimento embriológico e fetal, alinhado com diagnósticos precoces e informações de fácil acesso, pode contribuir no preparo social e psicológico dos familiares, principalmente da saúde materna, o que foi fortemente ressaltado em estudo realizado por Pereira et al. (2011), em que os autores evidenciam que mais pesquisas seriam necessárias para mensurar os transtornos familiares em decorrência de indivíduos com ACs na família.

No estudo de Santos e Dias (2005) há a seguinte frase: “*Os defeitos congênitos podem ser consideravelmente reduzidos através da educação em saúde*”. Na maioria dos casos, a informação adequada pode permitir o aconselhamento genético adequado (esclarecendo a causa) e, assim, atuar no sentido da prevenção primária.

Segundo Castilha e Orioli (2000 apud SANTOS; DIAS, 2005), atualmente existem métodos que podem prevenir cerca de 70% das anomalias congênitas, e as demais necessitam de estudos e pesquisas. Dentre estas prevenções, há três tipos de mecanismos a serem considerados, sendo eles: *primária*, que ocorre no período pré-concepcional, quando o tratamento ocorre em pessoas saudáveis, evitando a doença na tentativa de diminuir a exposição aos fatores de riscos; *secundária*, ocorre no período pré-natal, atuando em pessoas com a manifestação da doença, ao evitar sua evolução e maiores sequelas através do diagnóstico precoce, propondo o tratamento correto; *terciária*, período mais frequente no pós-natal, atuando em pessoas que possuem a doença, na tentativa de minimizar os danos à saúde através da reabilitação e adaptação.

No Brasil, há evidências de que exista um maior número de prestação de serviços de genética nas capitais comparado às cidades do interior, o que contribui nas dificuldades dos

aspectos de referência e contrarreferência destes indivíduos no seu acompanhamento (BRITO et al., 2010).

Atualmente existe um programa de vigilância epidemiológica presente no Brasil e na América Latina, cuja sede fica no departamento de genética do Instituto Oswaldo Cruz, no Rio de Janeiro, conhecido por Estudo Colaborativo Latino - Americano de Malformações Congênitas (ECLAMC), que foi criado na tentativa de reter registros das ACs em recém-nascidos através de estudos epidemiológicos com base hospitalar. Com finalidade de promover políticas de saúde, este programa conta com 70 hospitais partilhados entre dez países da América do Sul, Costa Rica e a República Dominicana (ECLAMC, 2016).

O desenvolvimento e implantação do sistema para a gestão da informação e vigilância de AC, proposto neste estudo, teve a iniciativa inédita no país de agrupar as informações públicas nacionais, cobrindo informações de nascimento, internação e desfecho do óbito para as notificações das anomalias e esta característica que a distingue e vem contribuir com as produções e pesquisas geradas pelo ECLAMC, que é referência nesta temática.

Sendo assim, tem-se a ciência que o Brasil conta com sistemas de informações nacionais, de cujos bancos públicos são facilmente extraídos dados sobre estas anomalias. Dentre estes sistemas, temos três que serão estudados neste trabalho: o Sistema de Informações Hospitalares do SUS (SIH), o Sistema de Informações sobre Nascidos Vivos (SINASC) e o Sistema de Informações sobre Mortalidade (SIM).

1.2. SISTEMAS DE INFORMAÇÃO NO BRASIL: SINASC, SIH E SIM

O Sistema de Informações sobre Nascidos Vivos tem como gestor o Centro Nacional de Epidemiologia (CENEPI), da Fundação Nacional de Saúde (FUNASA). O documento preenchido e padronizado nacionalmente de uso deste sistema é a Declaração de Nascido Vivo (DN), sendo aprovado o seu uso a partir de janeiro de 1990 (BRASIL, 2001a). O principal intuito do Ministério da Saúde em investir neste sistema se deve ao fato de este ser importante na produção de diversas análises estatísticas sobre as condições dos nascidos vivos e da saúde materna. Sua implantação ocorreu gradualmente e de maneira lenta ao longo do território nacional. Este corresponde ao primeiro sistema em informações de estatísticas vitais desenvolvidos computacionalmente.

Há um fluxo geral que a Declaração de Nascido Vivo percorre. A princípio ela é fornecida gratuitamente às secretarias estaduais de saúde que, as encaminham às secretarias

municipais de saúde, essas por sua vez, distribuem aos cartórios e estabelecimentos de saúde locais. Este fluxo é sugerido pelo Ministério da Saúde e pode ser visto no anexo A. O fluxo da informação da DN em que os partos ocorreram nos hospitais ou domiciliares podem ser vistos nos anexos B, Ca e Cb, respectivamente.

A DN possui preenchimento obrigatório nacionalmente para todos os nascidos vivos. A informatização da DN é de responsabilidade municipal quando este tiver o sistema informatizado. Caso não haja, o município deve passar os dados à instância estadual, que por sua vez centralizará os dados e enviará à instância federal, que terá a responsabilidade de disponibilizar as informações publicamente.

De interesse deste trabalho em particular, o campo 34 da DN foi criado neste sistema no ano de 2000 para permitir a obtenção de informações sobre a anomalia congênita de acordo com a Classificação Internacional de Doenças (CID10) capítulo XVII – “Malformações Congênitas, Deformidades e Anomalias Cromossômicas” (BRASIL, 2016). Este campo é composto por duas variáveis, sendo que uma identifica a presença ou ausência de AC e a outra variável possibilita visualizar seu código (CID10). O que se nota nas informações disponibilizadas pelo DATASUS é que no ano de 2000 somente a variável identificadora é apresentada e a partir de 2001 as duas compõem o banco de dados. Este fato é relevante já que neste projeto utilizamos as duas variáveis e, portanto, nosso estudo se inicia a partir de 2001. No ano de 2011, a DN passou por algumas modificações, dentre elas, o campo 34 foi reestruturado, e a variável que identifica a AC passa a constar no Bloco I, campo 6, e a variável com a informação do CID passa a se localizar no bloco VI, campo 4. No entanto, para o projeto esta mudança não afeta os cálculos e estudos (BRASIL, 2005).

Assim como o SINASC, o Sistema de Informações Hospitalares do Sistema Único de Saúde também foi criado pelo Ministério da Saúde no ano de 1990, sendo incumbência do Instituto Nacional de Assistência Médica e Previdência Social (INAMPS) em efetuar sua implantação em nível nacional (BRASIL, 2004a).

O SIH possui informações que tipicamente são utilizadas para efetuar o pagamento dos serviços hospitalares. Todos os hospitais participantes do SUS, sejam eles privados, sejam públicos, necessitam enviar informações referentes às internações, por meio do documento padrão estabelecido e conhecido por Autorização de Internação Hospitalar (AIH). Estes dados são enviados aos gestores municipais ou estaduais. O fluxo do sistema pode ser visto no anexo D.

A verificação de notificação de informações referente à anomalia congênita no SIH foi através de duas variáveis denominadas de diagnóstico principal e diagnóstico secundário, que

também são representadas pelo CID 10, do capítulo XVII, o que também as torna uma fonte de dados importante para o nosso estudo (BRASIL, 2015).

O Sistema de Informações sobre Mortalidade possui o CENEPI como gestor. O documento padrão utilizado nacionalmente é a Declaração de Óbito (DO), que também é de preenchimento obrigatório. Este sistema tem por intuito captar dados de mortalidade em todas as instâncias e entidades de saúde. O SIM foi criado pelo Ministério da Saúde em 1975 com a finalidade de se ter um sistema de vigilância epidemiológica no país. Foi nos primeiros anos da década de 1990, através de técnicas estabelecidas pela FUNASA, que os dados passaram a ser digitalizados (BRASIL, 2001b).

A DO possui um fluxo geral que segue a seguinte distribuição: a princípio ela é fornecida às secretarias estaduais de saúde, que as encaminham às secretarias municipais de saúde. Essas, por sua vez, repassam aos estabelecimentos de saúde, institutos médico-legais, serviços de verificação de óbitos, cartórios de registro civil e médicos, conhecidos como Unidades Notificadoras (Anexo E). Nestes locais há a coleta dos dados que, ao serem efetuados, seguirão o fluxo inverso para devolução da informação às secretarias municipais, onde será realizado o primeiro processamento dos dados e encaminhados às secretarias estaduais, que por sua vez transmitirão ao CENEPI, responsável por gerir os dados nacionalmente, fazendo a consolidação das bases de dados sobre mortalidade, disponibilizando-as publicamente. Os fluxos da informação de óbitos hospitalares e por causas naturais sem a presença de um médico podem ser vistos nos anexos F e G.

Para este projeto, o SIM conta com o campo 49, que é composto pelas seguintes variáveis: linha A, linha B, linha C, linha D, linha II e causa básica, que são responsáveis por conter informações referentes ao CID10. Em especial consideramos as informações referentes ao capítulo XVII para efetuarmos estudos e análises relevantes ao nosso tema central AC (BRASIL, 2006).

Consequentemente, estes sistemas SINASC, SIH e SIM são potencialmente excelentes fontes de dados por conter registros com informações referentes às ACs, a cada ano, em todo o território nacional. As DNs e DOs são de preenchimento obrigatório, o que nos possibilita uma cobertura quase que total destas informações. Já as AIHs somente nos trazem informações para aqueles indivíduos que acessam o serviço hospitalar do SUS.

Atualmente os dados públicos dos três sistemas estão disponibilizados na internet pelo Departamento de Informática do SUS (DATASUS), sendo que a sua implantação ocorreu de forma lenta e gradual em todos os estados do país (BRASIL, 2017c).

1.3. MOTIVAÇÃO E JUSTIFICATIVA

Acredita-se que este projeto seja estruturante da atenção à saúde em Genética Clínica, minimize desigualdades regionais e agregue informações epidemiológicas, na medida em que o produto principal foi o desenvolvimento de um sistema que fornece informações sobre a notificação de casos de anomalias congênitas e que permite uma avaliação do trânsito de um agravo específico dentro do sistema de saúde. Isto, disponibilizado via web, possibilita que diversos grupos pelo Brasil possam avaliar regionalmente e nacionalmente estas questões.

2. OBJETIVO

2.1. GERAL

Desenvolver sistema de vigilância de anomalias congênitas utilizando como base bancos de dados disponíveis através do DATASUS (SINASC, SIH e SIM). Com esse sistema tem-se um melhor aproveitamento das informações disponíveis, mapeando os nascidos vivos, as internações e os desfechos do óbito pela busca de CIDs ligados às anomalias congênitas.

2.2. SECUNDÁRIO

- Automatizar tanto estudos retrospectivos bem como contribuir para a formação de uma base de dados nacional de anomalias congênitas, com uma infraestrutura permanente para a base de dados relacionada através da técnica do linkage probabilístico.
- Implementar um sistema (software) com módulos para a gestão da informação, através de módulos de auditoria de dados e de análise da epidemiologia das anomalias congênitas, capaz de mostrar estatísticas regionalizadas tanto da incidência ou prevalência, como da assistência a esse agravo.
- Estudar e estabelecer dentro do sistema a capacidade de rastreabilidade de pacientes que ao nascerem com AC sofreram internações ao longo dos anos devido a um agravo específico e/ou faleceram.

3. MATERIAIS E MÉTODOS

3.1. TIPO DE ESTUDO

O modelo de estudo adotado neste projeto corresponde ao descritivo, sendo este retrospectivo por abranger o período de janeiro/2001 a dezembro/2011.

A pergunta central do estudo é: “O sistema permite acompanhar o processo de vigilância e monitoramento dos indivíduos notificados com anomalias congênitas nas bases de dados do SINASC, SIH e SIM, descrevendo os impactos e disponibilizando as informações públicas através de ferramentas de buscas epidemiológicas?”.

3.2. FONTE DE DADOS

A população de referência corresponde aos dados de todo o território brasileiro oriundos dos sistemas públicos nacionais SINASC, SIH e SIM, do período de janeiro/2001 a dezembro/2011, disponibilizados publicamente no site do DATASUS.

A população de estudo é representada pelos registros que contêm as informações dos indivíduos acometidos por AC de todos os estados do território brasileiro.

Adotamos como critério de inclusão dos dados todos os registros disponibilizados pelo DATASUS dos sistemas SINASC e SIM. No SIH teremos apenas os registros com CID referente ao capítulo XVII nos campos “diagnóstico principal” ou “diagnóstico secundário”. Já os critérios de exclusão foram compostos pelos registros que não se adequaram aos critérios de inclusão, no caso do SIH.

3.3. BANCO DE DADOS

O software que utilizamos para compor o banco de dados corresponde ao PostgreSQL que é um projeto de Sistema Gerenciador de Banco de Dados (SGBD) completamente gratuito e de código aberto. Foi criado em 1986 na cidade da Califórnia, na Universidade de Berkeley.

Por estar há tanto tempo ativo, sua arquitetura é considerada confiável e tem como característica manter a integridade dos dados. Este SGBD é compatível com a maioria dos

sistemas operacionais conhecidos e possui suporte a chaves estrangeiras, junções, visões, gatilhos e procedimentos (em diferentes linguagens), características estas que garantem a qualidade deste software. Ele também possui funcionalidades sofisticadas como: controle de concorrência multiversionado, recuperação em um ponto de tempo, replicação assíncrona, transações agrupadas, cópias de segurança, planejador de consultas e transações sequencial para tolerância a falhas (THE POSTGRESQL GLOBAL DEVELOPMENT GROUP, 2017). Por estas razões este SGBD foi escolhido.

A princípio foram extraídos os dados referentes ao SINASC e SIH do site DATASUS na extensão compactada *.dbc* e que foram expandidos para o formato *.dbf* através do tabulador denominado *tabwin32* (programa também disponibilizado gratuitamente pelo site do DATASUS). A partir da expansão, os dados foram passados para arquivos *.sql* que foram adicionados ao banco de dados. Posteriormente foi realizado o mesmo processo nos dados do SIM.

O DATASUS disponibiliza arquivos anualmente por estado para as bases SINASC e SIM, desta forma foi criadas tabelas organizadas por estado agregando todas as informações existentes no site. Já o SIH foi construído apenas com dados de registros que apresentavam o campo diagnóstico principal ou secundário com CID referente ao capítulo XVII, através de filtragem realizada no *tabwin32*. Para esta base foi criada uma única tabela nacional, e por esta razão o denominaremos a partir de agora de SIHAC.

Com o banco estruturado, iniciou-se o desenvolvimento do portal.

3.4. DESENVOLVIMENTO DO PORTAL

O desenvolvimento e implantação do sistema para a gestão da informação e vigilância de anomalias congênitas no Brasil tornou-se o “Portal de Monitoramento de Anomalia Congênita - (PMAC)”, que foi inicialmente projetado para conter informações de AC respeitando as características de cada sistema disponibilizado pelo DATASUS. Desta forma, optou-se por criar menus (módulos) e submenus (ferramentas) separados por base de dados.

As informações no site estão distribuídas conforme apresentado na Figura 1 e descritos a seguir:

- a) **Início:** que apresenta ao usuário o Portal de Monitoramento de Anomalias Congênitas;

- b) **Histórico:** que permite o acesso a informações gerais sobre AC e sobre os sistemas nacionais SINASC, SIH e SIM;
- c) **Ferramenta SIH:** que permite a manipulação dos dados sobre as internações dos indivíduos com AC;
- d) **Ferramenta SINASC:** que permite a manipulação dos dados sobre nascidos vivos com AC;
- e) **Ferramenta SIM:** que permite a manipulação dos dados sobre óbitos de indivíduos acometidos por AC;
- f) **Links:** que contém acesso a referências e;
- g) **Quem Somos:** que apresenta os envolvidos no trabalho, a possibilidade de contato por e-mail e pesquisa por satisfação do usuário.

Figura 1 – Representação dos menus e submenus presentes no portal PMAC

3.4.1. REQUISITOS DO PORTAL

Análises de requisitos são de extrema importância na área de engenharia de software e podem garantir o sucesso do produto final resultante do sistema, pois é nesta etapa que se reúnem informações das necessidades do sistema e de quem as dela faz uso. Estes requisitos são divididos em funcionais e não funcionais (PRESSMAN, 2006).

Os requisitos funcionais descrevem as funcionalidades desejadas pelo sistema. As que foram estabelecidas no desenvolvimento do portal são:

- O portal deve apresentar menu interativo quando o usuário acessar o site. Este menu deve permitir o acesso aos submenus;
- O menu Inicio deve apresentar o portal e auxiliar o usuário na navegação;
- O menu Histórico deve permitir o acesso às informações relevantes do portal, descrever anomalia congênita, as bases SINASC, SIH, SIM e sobre o linkage probabilístico;
- O menu Ferramenta SIH deve abrir submenus que permitam buscar dados das internações e analisar as AIHs;
- O menu Ferramenta SINASC deve abrir submenus que permitam buscar dados dos nascidos vivos e analisar as DN;
- O menu Ferramenta SIM deve abrir submenus que permitam buscar dados de mortalidade e analisar as DO;
- O menu Links deve permitir ao usuário pesquisar mais informações das fontes originais referenciadas no site;
- O menu Quem Somos deve permitir ao usuário conhecer quem são os responsáveis pelo site, como entrar em contato e opinarem através de uma pesquisa de satisfação;
- Dentro das ferramentas deve haver filtros referenciando o nível de localização da pesquisa, se esta será estadual e/ou municipal. Quando houver restrições, avisos ao usuário devem aparecer;
- Nas ferramentas que usam as informações de anomalia congênita (casos específicos) deve haver filtros de pesquisa nos três níveis de hierarquia do CID: por grupo, categoria e subcategoria;
- Todas as ferramentas devem conter uma breve explicação, antes de apresentarem os filtros de buscas;

- As ferramentas devem apresentar seus resultados em formato de fórmula, tabelas ou gráficos;
- Nos resultados, o portal deve permitir retornar a página com os filtros ou navegar em outros menu/submenus;
- O portal deve permitir realizar novas buscas.

Os requisitos não funcionais descrevem como o sistema deve se comportar. As funcionalidades que foram estabelecidas no desenvolvimento do portal são:

- Usabilidade: o portal deverá apresentar interface agradável, simples, acessível e intuitiva mantendo padrões entre as telas;
- Eficiência: deve garantir um bom desempenho na realização de suas tarefas, sem alterar a consistência dos dados trazidos nas buscas pelos filtros;
- Confiabilidade: o portal deverá garantir a integridade dos dados apresentados pelo usuário;
- Manutenibilidade: o portal deverá ser fácil em efetuar melhorias ou correções.

3.4.2. RECURSOS DE DESENVOLVIMENTO DAS INTERFACES GRÁFICAS

Na área da computação, o desenvolvimento das interfaces gráficas e das funcionalidades do portal recebe termos generalizados de frontend e backend. O frontend coleta os dados e informações inseridos através da interação do usuário e os envia para o backend executar no servidor web, pois este é responsável por implementar as regras de negócio.

Neste projeto, o frontend é composto pelas linguagens de marcação de hipertexto, do inglês HyperText Markup Language (HTML), a formatação do Cascading Style Sheets (CSS) e a execução das ações e comportamentos do JavaScript. No backend tem-se o Hypertext Preprocessor (PHP), do qual o código PHP fica embutido junto ao do HTML.

O HTML é um recurso simples de ser usado e acessível para a produção de documentos. Sua codificação permite definir os componentes da página, separar o cabeçalho, corpo do texto e rodapé, bem como agregar imagens e vídeos. A codificação pode ocorrer em qualquer editor de texto comum, como o bloco de notas, e o seu resultado é facilmente interpretado por diferentes navegadores web existentes. Esta linguagem nos auxiliou principalmente na formatação e padronização da estrutura com a produção e exibição de

textos informativos, tabelas, imagens e gráfico (THE WORLD WIDE WEB CONSORTIUM, 2017).

Para complementar o HTML há o CSS. Cujas criação desta linguagem se deu no intuito de auxiliar nas definições do estilo e layout dos textos nas páginas web, ou seja, definir a aparência de como as informações são abordadas, tudo através da possibilidade de inserir cores, alterar formatos de fontes, espaçamentos, alinhamento, dentre outras funcionalidades de layout (W3SCHOOLS, 2017a).

O terceiro recurso usado pelo frontend foi o JavaScript (W3SCHOOLS, 2017b), que tem por finalidade programar o comportamento das páginas e controlar as ações e funções definidas pelo HTML e CSS.

O PHP é uma linguagem que pode ser usada em diferentes sistemas operacionais e tem suporte para a maioria dos servidores web existente. Esta característica é importante uma vez que seu código é lido e interpretado pelo servidor e não na máquina cliente (do usuário). Uma das principais características desta linguagem é permitir dinamicidade às páginas criadas, tornando-as rápida. Também oferece suporte à coleta de dados de formulários, gera páginas com conteúdo dinâmico e/ou envia e recebe mensagens, suporta acesso a banco de dados e muito mais (THE PHP GROUP, 2017).

A combinação de todos estes recursos torna possível o acesso e a execução das funcionalidades do portal PMAC. Vale lembrar que estas linguagens são independentes umas das outras, ou seja, outras linguagens poderiam ter sido adotadas para desenvolver o site, no entanto, pensando na manutenibilidade, foram selecionadas estas linguagens por serem de ampla difusão e conhecimento dos desenvolvedores web.

3.4.3. ETAPA DE TESTE E AVALIAÇÃO

Ao final do desenvolvimento, iniciou-se a etapa de teste permitindo avaliar o conjunto de requisitos e funcionalidades do portal. Neste projeto foi empregado o teste caixa-preta junto com a técnica de avaliação heurística.

Para analisar os requisitos funcionais do sistema através da verificação da usabilidade do portal, foi utilizado o teste conhecido como caixa-preta. Neste teste o usuário desconhece os códigos internos e, usando as funções do sistema e os requisitos estabelecidos, consegue identificar problemas e/ou erros (PRESSMAN, 2006). Após a realização do teste, seguiu-se para a etapa da avaliação.

A avaliação heurística é uma técnica usada na área de avaliações de interação usuário computador, em inglês Human Computer Interaction (HCI), cujas principais características são: facilidade em ensinar a técnica, rapidez na execução e respostas, custo baixo e grupo pequeno de avaliadores (03 a 05 integrantes). Foi criada por Jakob Nielsen e Molich em 1990, atribuindo dez itens a serem avaliados (NIELSEN, 2016), sendo eles:

- 1) **Visibilidade do estado do sistema:** o sistema deve sempre manter os usuários informados sobre o que está acontecendo, através de feedback apropriado, em um tempo de resposta razoável;
- 2) **Correspondência entre o sistema e o mundo real:** o sistema deve falar a linguagem do usuário, com palavras, frases e conceitos familiares, em vez de termos orientados ao sistema. Devem ser seguidas convenções do mundo real, fazendo com que as informações apareçam em uma ordem natural e lógica;
- 3) **Controle e liberdade do usuário:** os usuários muitas vezes escolhem as funções do sistema por engano e precisarão de uma "saída de emergência" claramente marcada para deixar o estado indesejado sem ter que passar por um diálogo estendido. Suporte para desfazer e refazer ações;
- 4) **Consistência e padrões:** os usuários não devem ter que se perguntar se diferentes palavras, situações ou ações significam a mesma coisa. Sistema deve exibir informações consistentes e padronizadas;
- 5) **Prevenção de erros:** melhor do que boas mensagens de erro é um sistema cuidadoso que impede que um problema ocorra em primeiro lugar. Elimine as condições propensas a erros ou procure por elas e apresente aos usuários uma opção de confirmação antes de se comprometerem com a ação;
- 6) **Reconhecimento em vez de lembrança:** minimiza a carga de memória do usuário, tornando visíveis objetos, ações e opções. O usuário não deve se lembrar de informações de uma parte do diálogo para outra. As instruções de utilização do sistema devem ser visíveis ou facilmente recuperáveis sempre que adequado;
- 7) **Flexibilidade e eficiência de uso:** aceleradores, invisíveis para usuário novo, mas muitas vezes podem acelerar a interação para o usuário especializado, de tal forma que o sistema pode servir tanto para usuários inexperientes, como experientes. Permite que os usuários adaptem ações frequentes;
- 8) **Design estético e minimalista:** os diálogos não devem conter informações irrelevantes ou raramente usadas. Cada unidade extra de informação em um

diálogo compete com as unidades de informação relevantes e diminui sua visibilidade relativa;

- 9) **Ajudar usuários a reconhecer, diagnosticar e se recuperar dos erros:** mensagem de erro deve ser expressa em linguagem simples (sem códigos). Indicar, com precisão, o problema e sugerir construtivamente uma solução;
- 10) **Suporte e documentação:** melhor seria o sistema não ter, mas pode ser necessário fornecer ajuda e documentação. Essas informações devem ser fáceis de pesquisar, centradas na tarefa do usuário, lista de passos concretos a serem realizados e não devem ser muito grandes.

Dentro destes dez itens são estabelecidos cinco critérios numa escala de 0 a 4 a serem atribuídos para medir a usabilidade do sistema, sendo eles: (0) - Não concordo que é problema de usabilidade; (1) - Problema cosmético: correção se tiver tempo; (2) - Problema menor: correção de baixa prioridade; (3) - Problema grave: correção de alta prioridade; (4) - Catástrofe de usabilidade: correção é imperativa.

3.5. AUDITORIA

Auditoria foi conceituada por Lambeck em 1956 como “avaliação da qualidade da atenção com base na observação direta, registro e história clínica do cliente” (BRASIL, 2004b).

Atualmente no país há o Sistema Nacional de Auditoria (SNA) do SUS, de responsabilidade do Departamento Nacional de Auditoria do SUS (DENASUS), que segue a definição proposta pelo ParticipaSUS, sendo: “A auditoria é um instrumento de gestão para fortalecer o Sistema Único de Saúde (SUS), contribuindo para a alocação e utilização adequada dos recursos, a garantia do acesso e a qualidade da atenção à saúde oferecida aos cidadãos.” (BRASIL, 2011).

A finalidade da auditoria neste trabalho consiste em verificar a qualidade das informações dos dados nos bancos públicos, ou seja, com o banco estruturado foi realizado inicialmente uma auditoria de maneira a encontrar discrepâncias entre as bases, bem como eventuais subnotificações e erros de preenchimento, as quais analisem a autenticidade e confiabilidade dos dados.

Esta verificação da qualidade dos campos pela auditoria garante o sucesso na escolha das variáveis a serem utilizadas na técnica do relacionamento probabilístico, permitindo assim o desenvolvimento de estudos de mapeamento dos casos registrados com AC.

Sendo assim, três análises foram realizadas e são explicadas a seguir.

3.5.1. EVOLUÇÃO DOS BANCOS E DOS DADOS

A primeira análise verificou a evolução da estrutura das tabelas dos bancos, ou seja, observou-se a quantidade de campos existentes, inseridos, alterados e deletados nas DNs, DOs e AIHs ao longo do período de janeiro/2001 a dezembro/2011.

Além das análises dos campos, obteve-se a evolução quantitativa de registros inseridos em cada base ao longo do período de estudo.

3.5.2. INCOMPLETUDE

A segunda análise consiste na incompletude que corresponde à taxa de campos (valores) em branco de um registro referente aos dados do SINASC, SIHAC e SIM, ou seja, uma informação incompleta ou não notificada de cada variável. Os resultados foram adquiridos através da seguinte fórmula aplicada pela ferramenta desenvolvida automaticamente:

$$\text{Incompletude} = \frac{\text{Quantidade de registros em branco}}{\text{Total de registros}} \times (100) \quad (1)$$

Os resultados gerados a partir desta análise foram classificados com os seguintes graus de avaliação conforme Figura 2 (ROMERO; CUNHA, 2002):

- a) **“Excelente”**, para as variáveis que apresentaram resultados menores que 5%;
- b) **“Bom”**, para as que estiveram no intervalo de 5% a 10%;
- c) **“Regular”**, para as que estiveram entre 10% e 20%;
- d) **“Ruim”**, para as que tiveram resultados entre 20% e 50% e;
- e) **“Péssimo”**, para as que apresentaram resultados acima de 50%.

Figura 2 – Representação da classificação segundo os graus de avaliação que vai de 0% a 100%

3.5.3. INCONSISTÊNCIA

A terceira análise realizada corresponde à inconsistência, que é a taxa de erro ou discordância de um determinado campo ou conjunto de campos.

Este projeto foi especificamente calculado sobre o campo 34 (período de 2001 a 2010) do SINASC, que abrange duas variáveis: (a) “idanomal”, que indica a presença, ausência ou ignorado de AC e; (b) “codanomal”, que especifica a AC de acordo com o CID10. No ano de 2011, a DN passou por algumas modificações, dentre elas, o campo 34 foi reestruturado e a variável “idanomal” passa a constar no Bloco I, campo 6, e o “codanomal” se localiza no bloco VI, campo 4. No entanto, para o projeto, esta mudança não afeta o cálculo da inconsistência nem as demais ferramentas.

A utilização desta análise nestes campos ocorre em razão de as variáveis serem correlacionadas e o preenchimento de uma interferir diretamente na outra. O intuito é verificar quando há um preenchimento onde não deveria acontecer, ou seja, quando as duas variáveis não concordam em relação às informações apresentadas. O cálculo da inconsistência é expresso pela Fórmula 2, em que ID corresponde à variável que identifica a presença (1), ausência (2) ou ignorado (9) de anomalia, e COD identifica o código da anomalia sendo preenchido (+) ou ausente (-).

$$\text{Inconsistência} = \frac{\text{ID}(1)\text{eCOD}(-) + \text{ID}(2)\text{eCOD}(+) + \text{ID}(9)\text{eCOD}(+)}{\text{ID}(1) + \text{ID}(2) + \text{ID}(9)} \times (10.000) \quad (2)$$

As variáveis que contêm as informações de AC no SIHAC são o diagnóstico principal e secundário, as do SIM são linhaa, linhab, linhac, linhad, linhaid, causa básica e causa básica original. Diferente das variáveis idanomal e codanomal, estas variáveis não são correlacionadas e por esta razão não realizou-se a análise de inconsistência no SIHAC e SIM.

3.6. COEFICIENTES EM EPIDEMIOLOGIA

Estudos de coeficientes em epidemiologia são utilizados para medir o nível de saúde ou doença de uma população (FRANCO; PASSOS, 2011).

O cálculo dos coeficientes pode ser definido como medida de proporção dos quais “*eventos do numerador representa um risco de ocorrência em relação ao denominador*” (MERCHÁN-HAMANN; TAUIL; COSTA, 2000).

Já a medida taxa vem sendo utilizada como “taxa de ocorrência de um dado evento incidente em termos da sua tendência em um período de tempo” (MERCHÁN-HAMANN; TAUIL; COSTA, 2000).

Estas conceituações são importantes no nosso estudo, uma vez que utilizamos medidas de morbidade e mortalidade com estas nomenclaturas no desenvolvimento das ferramentas do portal. Veremos a seguir os conceitos e cálculos utilizados e implementados no site.

3.6.1. COEFICIENTE DE INCIDÊNCIA

Incidência corresponde a um número de novos casos de uma determinada doença num período de tempo estudado, tendo como base uma população exposta a esta doença.

Já o coeficiente de Incidência seria a “proporção de casos novos de uma doença na população que esteve exposta ao risco de adoecer, durante o período considerado” (FRANCO; PASSOS, 2011).

$$\text{Coeficiente de Incidência} = \frac{\text{Número de novos casos da doença num período de tempo em uma população}}{\text{Número de pessoas sob o risco de desenvolver a doença durante o período de tempo}} \times (\text{base}) \quad (3)$$

A equação vista acima foi retirada da referência de Franco e Passos (2011), e a base pode ser substituída por valores como 100, 1.000, 10.000 ou 100.000.

Houve uma discussão se adotaríamos a medida incidência ou prevalência para o cálculo dos registros do Sistema de Informação sobre Nascidos Vivos. Esta discussão baseia-se na premissa de que crianças que nasceram com AC tiveram a incidência deste agravo no período gestacional e por esta razão alguns geneticistas defendem a ideia do conceito de prevalência de AC. No entanto, neste projeto utilizamos como padrão o termo incidência de notificação de AC no SINASC, justificando desta forma nossa escolha de medida de morbidade.

Sendo assim, na Fórmula 3 foi substituído o número de novos casos da doença por número de notificação de nascimento com AC, adotou-se a base com valor 10.000 no período de estudo de janeiro/2001 a dezembro/2011, tendo como resultado a Fórmula 4 vista a seguir:

$$\text{Coeficiente de Incidência} = \frac{\text{Número de notificação de nascimentos com AC de 2001 a 2011}}{\text{Número total de nascimentos de 2001 a 2011}} \times (10.000) \quad (4)$$

3.6.2. COEFICIENTE DE MORTALIDADE INFANTIL

O coeficiente de mortalidade infantil, também conhecido como taxa de mortalidade infantil, corresponde ao número de crianças menores de um ano que faleceram para cada 1.000 nascidos vivos numa determinada população e período, como visto na Fórmula 5.

$$\text{Coeficiente de Mortalidade Infantil} = \frac{\text{Número de óbitos de crianças menores de um ano}}{\text{Número de nascidos vivos}} \times (1.000) \quad (5)$$

Neste projeto o numerador da Fórmula 5 foi representado pelo número de crianças menores de um ano que vieram a óbito no ano de estudo. O período compreende janeiro/2001 a dezembro/2011 e no desenvolvimento da ferramenta permite-se selecionar apenas um ano de cada vez dentro do período de estudo adotado.

3.6.3. TAXA DE MORTALIDADE PERINATAL

A taxa de mortalidade perinatal corresponde ao número de óbitos fetais a partir de 22 semanas completas de gestação, acrescido dos óbitos ocorridos até o sétimo dia de vida pelo

número de nascidos vivos e óbitos fetais numa determinada população e período, como visto na Fórmula 6.

$$\text{Taxa de Mortalidade Perinatal} = \frac{\text{Óbitos fetais a partir de 22 semanas de gestação + crianças com até 7 dias de vida}}{\text{Numero de Nascidos Vivos e Óbitos fetais}} \times (1.000) \quad (6)$$

Uma observação retratada pelo DATASUS deve ser levada em consideração no cálculo da Fórmula 6, sendo ela: “Considerando a subnotificação de óbitos fetais e a precariedade da informação disponível sobre o tempo de gestação, recomenda-se que, no cálculo do indicador, seja utilizado o número total de óbitos fetais informados, independentemente da duração da gestação” (BRASIL, 2017b).

3.6.4. TAXA DE MORTALIDADE NEONATAL TARDIA

A taxa de mortalidade neonatal tardia leva em conta o número de óbitos de 7 a 27 dias de vida completos pelo número de nascidos vivos residentes no local e ano selecionado, como visto na Fórmula 7.

$$\text{Taxa de Mortalidade Neonatal Tardia} = \frac{\text{Óbitos de 7 a 27 dias de vida completos}}{\text{Numero de Nascidos Vivos}} \times (1.000) \quad (7)$$

A fórmula acima é o cálculo direto que foi utilizado no desenvolvimento da ferramenta no portal. No entanto, no site do DATASUS há a informação de que se pode calcular através de outra alternativa, sendo ela descrita como: “Aplica-se, à taxa de mortalidade infantil estimada pelo IBGE, a proporção de óbitos de 7 a 27 dias de vida completos informados no SIM (percentual em relação ao total de óbitos de menores de um ano, excluídos os de idade ignorada)” (BRASIL, 2017a).

3.6.5. MORTALIDADE SEGUNDO CAUSAS

Segundo Franco e Passos, pode-se utilizar taxas ou proporções para medir as causas de um determinado capítulo da CID-10, por grupo de doença ou por uma doença específica.

Sendo assim, utilizaremos este recurso no capítulo XVII para verificarmos a AC. Sua fórmula está expressa logo abaixo:

$$\text{Taxa de Mortalidade segundo causa} = \frac{\text{Total de óbitos por AC numa área e período escolhido}}{\text{População desta área e período escolhido}} \times (1.000) \quad (8)$$

3.7. RELACIONAMENTO PROBABILÍSTICO

A ideia, neste projeto, de relacionar os dados contidos nos diferentes bancos se baseia na premissa de encontrar uma ligação entre os registros que permita informar e/ou monitorar um indivíduo com AC nos diferentes sistemas públicos, ou seja, depois de notificado o nascimento buscou-se evidências de ligação entre os registros de um indivíduo, verificando se este gerou internações e/ou teve como desfecho o óbito. Atualmente existem dois métodos que realizam o relacionamento. Para escolher qual deles melhor se adapta aos dados, temos que analisar os campos presentes nas bases, além de conhecer as diferenças entre estes métodos.

O primeiro método chama-se determinístico e é utilizado quando o banco de dados apresenta um campo com identificador unívoco (por exemplo: o CPF, RG ou um número de identificação). Na ausência deste tipo de campo há a opção de usar o segundo método conhecido por probabilístico. Este método tem por objetivo reunir um conjunto de campos que sejam comuns entre os bancos de dados e verificar a probabilidade de os registros serem de um mesmo indivíduo (COELI; CAMARGO, 2002).

Os três bancos de dados utilizados neste projeto não apresentam campo unívoco, desta forma adotou-se o método probabilístico. Para efetuar este relacionamento utilizou-se o software livre chamado de OpenRecLink (CAMARGO; COELI, 2016). Alguns passos sugeridos por Coeli e Camargo (2007), foram realizados para obter o resultado fornecido pela técnica, sendo eles: a padronização dos campos comuns entre os bancos de dados; a blocagem, criando blocos lógicos para comparar os arquivos e deixar o processamento mais rápido (Figura 3); o pareamento, que é a utilização de algoritmos que comparam os campos; o cálculo do score, que revela quanto que os registros concordam entre si; definir os limiares superiores e inferiores, para classificar os registros em pares verdadeiros, pares falsos e pares duvidosos; fazer a verificação dos pares duvidosos manualmente.

O relacionamento probabilístico foi realizado a princípio entre os bancos SINASC e SIHAC. Posteriormente entre SINASC e SIM. As etapas apresentadas na sequência levam em consideração estes dois linkage.

Figura 3 – Representação ilustrativa das etapas da padronização e relacionamento

3.7.1. PADRONIZAÇÃO

Esta foi a primeira etapa realizada antes do relacionamento probabilístico e é um dos processos mais importantes, uma vez que se pode minimizar a possibilidade de erros.

A padronização dos dados tem por finalidade manter o formato dos campos e das variáveis nos bancos de dados iguais em todos os arquivos, propiciando assim o relacionamento entre eles (por exemplo: manter o formato da variável data ou a mesma quantidade de caracteres, etc.). Também permite selecionar apenas os campos que são relevantes no estudo com a remoção daqueles que não são utilizados, tornando assim este arquivo mais rápido na sua execução devido à redução no processamento, mantém somente os campos-chave usados nos processos de bloqueio e pareamento, além dos que podem auxiliar na tomada da decisão sobre os possíveis pares serem falsos ou verdadeiros (COELI; CAMARGO, 2007).

Durante o processo de padronização, todos os campos foram convertidos para campos caracteres e adotou-se a opção “Elimina Pontuação”, retirando todos os sinais de pontuação. Nas variáveis que correspondem a datas, foram criadas novas variáveis separando os anos e os meses, para auxiliar nas próximas etapas do relacionamento.

Sendo assim, são apresentadas a seguir as variáveis com a padronização adotada:

3.7.1.1. VARIÁVEIS COM CAMPOS EM COMUM ENTRE OS SISTEMAS

As variáveis comuns e que existem em todas as bases de dados correspondem a: sexo, data de nascimento, código do município residente, código da anomalia e número de identificação do estado.

A primeira variável citada anteriormente e que corresponde ao “sexo” se encontra no SINASC e SIM com o valor 1 representando o masculino e 2 o feminino. No SIHAC o valor 1 também consiste em ser masculino, no entanto o valor 3 é utilizado para caracterizar o feminino, sendo assim foram padronizados os arquivos seguindo a correspondência do SINASC e SIM (1 é o masculino e 2 o feminino) (Tabela 1).

Tabela 1 – Padronização da variável “sexo”

Variável	SINASC	SIM	SIHAC	Padronizado	Legenda
Sexo	1	1	1	1	Masculino
	2	2	3	2	Feminino

A segunda variável consiste na data de nascimento. Além desta variável, todas as datas importantes presentes no estudo receberam o mesmo critério de padronização e, por esta razão, estão na tabela a seguir. No SINASC e SIM a formatação original corresponde a dia, mês e ano. No SIHAC é representado por ano, mês e dia. Com a padronização os arquivos ficaram com a formatação do SIHAC (aaaammdd / ano mês e dia) (Tabela 2).

Tabela 2 – Padronização das variáveis “data”

Variável	SINASC	SIM	SIHAC	Padronizado
Data Nascimento	ddmmaaaa	ddmmaaaa	aaaammdd	aaaammdd
Data Internação			aaaammdd	aaaammdd
Data Saída			aaaammdd	aaaammdd
Data do Óbito		ddmmaaaa		aaaammdd

A terceira variável corresponde aos códigos de município e estabeleceu-se como padrão manter 6 caracteres. Este campo passou por modificações com acréscimo de caracteres nos últimos anos. Nos primeiros anos de estudo, o tamanho dos campos e seus valores atribuídos correspondiam a seis caracteres (C6) e este tamanho já é suficiente para identificar a maioria dos municípios com abrangência do período de estudo adotado e com base nas informações que há disponível (Tabela 3).

Tabela 3 – Padronização da variável “código município”

Variável	SINASC	SIM	SIHAC	Padronizado
Código município	C6	C6	C6	C6

A quarta variável citada corresponde aos códigos das anomalias congênitas segundo o CID-10 (capítulo VXII). Em todos os bancos de dados, estas variáveis são preenchidas na maioria das vezes com 4 caracteres (C4). Em poucos casos encontra-se valores com três caracteres ou mais. Por esta razão a padronização adotada corresponde a quatro caracteres (Tabela 4).

Tabela 4 – Padronização das variáveis com valor “CID-10 (cap. XVII)”

Variável	SINASC	SIM	SIHAC	Padronizado
Código das anomalias	C4			C4
Diagnóstico Principal			C4	C4
Diagnóstico Secundário			C4	C4
Linha A		C4		C4
Linha B		C4		C4
Linha C		C4		C4
Linha D		C4		C4
Linha II		C4		C4
Causa Básica		C4		C4
Causa Básica Original		C4		C4

A última variável citada e comum entre os bancos de dados corresponde ao número que representa o estado, sendo que no SINASC e SIM ela é representada por 2 caracteres. No SIHAC, esta variável apresenta 6 caracteres. Na padronização as variáveis ficaram com a padronização do SINASC e SIM com dois caracteres. Ressaltamos que a remoção dos demais caracteres do SIHAC não prejudica a informação, pois os dois primeiros caracteres já identificam o estado (Tabela 5).

Tabela 5 – Padronização das variáveis “código do estado”.

Variável	SINASC	SIM	SIHAC	Padronizado
UF	C2	C2	C6	C2

3.7.1.2. VARIÁVEIS COM CAMPOS DIFERENTES ENTRE OS SISTEMAS

As variáveis que são diferentes nos bancos de dados e que são relevantes para o estudo são utilizadas para auxiliar a revisão manualmente quando os resultados derem “pares duvidosos”. Estas variáveis não tiveram nenhuma alteração de padronização. Elas apenas foram incluídas nos arquivos.

No SINASC temos a variável-chave que contém o número da DN, a variável que identifica se o paciente contém uma anomalia, o código de estabelecimento de saúde, tipo de parto, hora do nascimento, raça/cor e a variável que corresponde ao local de nascimento.

No SIM temos a variável-chave que contém o número da DO, o tipo do óbito, hora do óbito, idade do paciente, local de ocorrência do óbito, código do estabelecimento, o código do município do óbito, morte em relação ao parto, morte durante a gravidez, morte durante o puerpério e morte no parto.

Já no SIHAC temos a variável-chave que contém o número da AIH, a variável morte e a variável que corresponde à idade do paciente.

3.7.2. RELACIONAMENTO

Como já mencionado, neste projeto adotaram-se duas etapas de relacionamento probabilístico. A primeira foi realizada entre o SINASC e SIHAC; e a segunda entre o SINASC e SIM.

Para realizar o relacionamento entre os arquivos precisou-se passar por dois processos, sendo eles: a blocagem e o pareamento.

Como visto anteriormente, a padronização dos dados auxiliou na etapa do relacionamento, pois propiciou que os campos (variáveis) comuns tivessem a mesma formatação, através da mesma quantidade de caracteres e não necessariamente o mesmo nome de variável.

3.7.2.1. *BLOCAGEM*

O processo de blocagem permite a criação de blocos lógicos entre os arquivos que são relacionados. Nesta etapa os cálculos se dão através da multiplicação entre os arquivos, o que propicia um alto custo no processamento (COELI; CAMARGO, 2007). Estes custos, sem a criação da blocagem nas bases relacionadas, podem ser vistos nas Tabelas 6 e 7.

Tabela 6 – Multiplicação entre arquivos relacionados SINASC-SIHAC.

	SINASCAC	SIHAC	Multiplicação
Acre	384	2.174	834.816
Alagoas	3.021	12.248	37.001.208
Amapá	313	1.956	612.228
Amazonas	5.234	8.185	42.840.290
Bahia	14.067	79.116	1.112.924.772
Ceará	8.528	33.636	286.847.808
Distrito Federal	2.647	20.845	55.176.715
Espírito Santo	3.031	13.488	40.882.128
Goiás	5.909	22.130	130.766.170
Maranhão	3.850	17.732	68.268.200
Minas Gerais	20.397	74.980	1.529.367.060
Mato Grosso	3.710	8.652	32.098.920
Mato G do Sul	2.496	8.245	20.579.520
Pará	7.656	20.252	155.049.312
Paraíba	6.267	12.246	76.745.682
Paraná	12.438	49.760	693.430.938
Pernambuco	13.351	37.786	504.480.886
Piauí	1.822	13.720	24.997.840
Rio de Janeiro	18.910	64.947	1.228.147.770
Rio G. do Norte	3.390	10.909	36.981.510
Rio G. do Sul	14.907	47.627	776.535.444
Rondônia	1.587	2.711	4.302.357
Roraima	525	1.315	690.375
Santa Catarina	8.051	26.735	215.243.485
São Paulo	56.928	267.681	15.238.543.968
Sergipe	2.688	5.395	14.501.760
Tocantins	1.522	4.102	6.243.244

Tabela 7 – Multiplicação entre arquivos relacionados SINASC-SIM.

	SINASC	SIM	Multiplicação
Acre	384	618	237.312
Alagoas	3.021	2.369	7.156.749
Amapá	313	559	174.967
Amazonas	5.234	3.361	17.591.474
Bahia	14.067	9.085	127.798.695
Ceará	8.528	6.367	54.297.776
Distrito Federal	2.647	2.356	6.236.332
Espírito Santo	3.031	2.885	8.744.435
Goiás	5.909	4.392	25.952.328
Maranhão	3.850	4.226	16.270.100
Minas Gerais	20.397	12.597	256.941.009
Mato Grosso	3.710	2.530	9.386.300
Mato G do Sul	2.496	2.205	5.503.680
Pará	7.656	5.338	40.867.728
Paraíba	6.267	2.640	16.544.880
Paraná	12.438	7.921	98.521.398
Pernambuco	13.351	7.466	99.678.566
Piauí	1.822	2.417	4.403.774
Rio de Janeiro	18.910	11.516	217.767.560
Rio G. do Norte	3.390	2.364	8.013.960
Rio G. do Sul	14.907	8.298	123.698.286
Rondônia	1.587	1.243	1.972.641
Roraima	525	468	245.700
Santa Catarina	8.051	4.151	33.419.701
São Paulo	56.928	31.069	1.768.696.032
Sergipe	2.688	1.883	5.061.504
Tocantins	1.522	1.036	1.576.792

Com a criação da blocagem, este processamento tornou-se otimizado, pois os blocos tiveram como prioridade dividir os registros. Esta técnica também aumentou a probabilidade de os pares serem verdadeiros, mas para que isto ocorresse precisou-se primeiramente criar a chave de blocagem que consiste na combinação de um conjunto de variáveis com o mesmo valor para esta chave preestabelecida (COELI; CAMARGO, 2007). Neste projeto, foram utilizadas as mesmas duas chaves para cada relacionamento (Tabelas 8), em que os registros não pareados na primeira chave foram comparados novamente com a segunda chave. A escolha das variáveis levou em consideração a diminuição do custo do processamento e a perda de pares verdadeiros, isto porque foram escolhidas variáveis com baixa probabilidade de ocorrência de erros.

Tabela 8 – Chaves de blocagem.

Passo	Chave
1	SEXO + MUNICÍPIO RESIDENTE + ANO NASCIMENTO
2	SEXO + ANO NASCIMENTO

3.7.2.2. PAREAMENTO

O processo de pareamento é uma das etapas do relacionamento que consiste na combinação entre os arquivos dos bancos de dados, cujas combinações apresentaram diferentes valores atribuídos aos escores, garantindo o quanto os pares são verdadeiros, falsos ou duvidosos (COELI; CAMARGO, 2007).

O valor final do escore foi construído a partir da ponderação da soma de todos os escores de cada variável utilizada no pareamento, pelo software OpenRecLink (COELI; CAMARGO, 2012), ou seja, todas as variáveis fizeram a sua contribuição de forma diferente no cálculo do escore.

Na composição do cálculo tem-se que, para cada variável representada pela letra i , há probabilidade de o par de registros concordarem. Este par poderá ser verdadeiro, representado na fórmula por m_i (verdadeiro positivo) ou será falso, representado na fórmula por u_i (falso positivo). Sendo que $1 - m_i$ corresponde à probabilidade da variável discordar, dado que se trata de um par verdadeiro (falso negativo) e que $1 - u_i$ representa a probabilidade de uma variável discordar, dado que se trata de um par falso (verdadeiro negativo) (COELI; CAMARGO, 2007).

Assim, dois fatores de concordância foram aplicados. O primeiro foi utilizado ao comparar dois arquivos e suas variáveis concordaram, denominado como fator de ponderação de concordância. Este fator é calculado como o logaritmo na base 2 da razão de verossimilhança entre m_i e u_i , como visto na formula a seguir:

$$(wc_i = \log_2 \left(\frac{m_i}{u_i} \right)), \quad (9)$$

O segundo fator foi utilizado quando dois registros foram comparados e suas variáveis não concordaram, denominado como fator de ponderação de discordância. O fator foi calculado como o logaritmo na base 2 da razão de verossimilhança entre $1 - m_i$ e $1 - u_i$, como visto a seguir:

$$(wd_i = \log_2 \left(\frac{(1 - m_i)}{(1 - u_i)} \right)), \quad (10)$$

O escore final foi obtido através da soma dos fatores de ponderação. A probabilidade m_i contribui positivamente no resultado já a probabilidade u_i contribui negativamente (COELI; CAMARGO, 2007).

Além dos fatores de ponderação, foi necessário utilizar algoritmos que possam efetuar a comparação das variáveis e retornar quanto elas são parecidas. A Tabela 9 e 10 apresentam as variáveis utilizadas no pareamento entre SINASC-SIM e SINASC-SIHAC respectivamente, assim como os algoritmos empregados em cada uma delas.

Nosso trabalho utilizou dois algoritmos, sendo eles: (1) Aproximado: Compara sequências de caracteres utilizando o algoritmo que calcula a distância de Levenshtein e retorna valores entre 1 (concordância total) e 0 (discordância total); (2) Caracter: Compara sequência de dígitos (ignorando separadores) comparando os pares de dígitos na mesma posição e retornando valores entre 1 (concordância total) e 0 (discordância total) (COELI; CAMARGO, 2007).

Tabela 9 – Variáveis utilizadas no pareamento SINASC - SIM com o algoritmo de comparação.

Variável	Algoritmo de Comparação	Concordância
Data do nascimento	Caracter	Total ou Parcial
Código da Anomalia	Aproximado	Total ou Parcial
Linha A	Aproximado	Total ou Parcial
Linha B	Aproximado	Total ou Parcial
Linha C	Aproximado	Total ou Parcial
Linha D	Aproximado	Total ou Parcial
Linha II	Aproximado	Total ou Parcial
Causa Básica	Aproximado	Total ou Parcial

Tabela 10 – Variáveis utilizadas no pareamento SINASC - SIHAC com o algoritmo de comparação.

Variável	Algoritmo de Comparação	Concordância
Data do nascimento	Caracter	Total ou Parcial
Código da Anomalia	Aproximado	Total ou Parcial
Diagnóstico Principal	Aproximado	Total ou Parcial

Uma vez definidas as chaves de bloqueio, os fatores de ponderação e os algoritmos de comparação, houve arquivos relacionados com seus respectivos escores. A última etapa do linkage correspondeu à classificação do par de registros, segundo os valores atribuídos pelo escore. Esta classificação foi em parte manual e parte automática. Neste último caso ocorreu devido à definição dos limiares superiores (mínimo aceitável) e inferiores (máximo recusado) que o próprio software gerou. Os escores que ficaram entre estes limiares são denominados como duvidosos e neste caso a análise foi feita manualmente.

3.8. ASPECTOS ÉTICOS

No desenvolvimento da pesquisa, foram cumpridas todas as exigências éticas prescritas pela resolução CNS 466/2012. Toda a equipe de pesquisadores esteve ciente e fez cumprir os conteúdos desta portaria. Vale ressaltar que a metodologia proposta não envolve experimentação em seres humanos, nem tem foco em grupos especiais. Como etapa preliminar, o projeto aqui apresentado foi encaminhado e aprovado pela Comissão de Ética do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto da USP, conforme o número de protocolo CAAE 45519915.7.0000.5440 e o número do parecer 1.120.062 (ANEXO H).

4. RESULTADOS

4.1. AUDITORIA

Como visto na seção 3.5, a auditoria verifica a qualidade das informações dos dados presentes nos bancos do SINASC, SIHAC e SIM. Os resultados de incompletude e inconsistência serão abordados após a apresentação da análise da evolução dos bancos e dos dados, que será vista a seguir.

4.1.1. EVOLUÇÃO DOS BANCOS E DOS DADOS

A verificação da evolução da estrutura das tabelas dos bancos é representada pela quantidade de campos existentes na DN, DO e AIH ao longo do período de janeiro/2001 a dezembro/2011. Alguns campos foram alterados, excluídos e outros criados.

O SINASC, SIM e SIH iniciaram 2001 com 23, 41 e 60 campos e, ao final de 2011, contavam com 59, 66 e 93 campos respectivamente. Estes resultados são apresentados na Figura 4.

Figura 4 – Representação da evolução da quantidade de campos nas bases de janeiro de 2001 a dezembro de 2011

Além das análises dos campos, foi realizada a verificação da evolução quantitativa de registros inseridos em cada base ao longo do período de estudo. Esta evolução foi abordada com o intuito de auxiliar na compreensão dos resultados obtidos nas ferramentas desenvolvidas. Para facilitar a leitura dos resultados, estes foram divididos por regiões do país.

A Figura 5 contém a representação dos dados do SINASC, em todas as regiões. Nota-se certa constância de inserção de registros dos estados ao longo dos anos. Através da média observa-se que os maiores valores inseridos são, por ordem decrescente, das regiões sudeste, sul, nordeste, centro-oeste e norte. O norte tem como exceção o estado do Pará, do qual seu resultado fica acima da média das regiões norte, centro-oeste e nordeste. Em contrapartida, na região sudeste tem-se que o estado do Espírito Santo fica abaixo das médias das regiões sudeste, sul e nordeste. A região sul conta com estados do Paraná e Rio Grande do Sul com uma leve diminuição dos registros ao longo do período estudado.

Figura 5 – Representação da evolução da quantidade de registros inseridos no SINASC de janeiro de 2001 a dezembro de 2011

O SIH é representado pela Figura 6 e assim como os resultados do SINASC uma característica das AIHs é que em todas as regiões nota-se certa constância de inserção dos registros nos estados de 2001 a 2011. As médias de preenchimento seguem em ordem decrescente: regiões sudeste, sul, nordeste, centro-oeste e norte. Na Bahia, estado do nordeste, verifica-se uma leve diminuição ao longo do período.

Figura 6 – Representação da evolução da quantidade de registros inseridos no SIH de janeiro de 2001 a dezembro de 2011

Já no SIM, que está representado pela Figura 7, há estados que se mantiveram constantes na inserção dos registros. No entanto, observa-se maior número de estados com leve aumento ao longo dos anos, mesma característica apresentada pela média que em algumas regiões se mantém constante e em outras cresce. Por ordem decrescente, tem-se: regiões sudeste, sul, nordeste, centro-oeste e norte.

Figura 7 – Representação da evolução da quantidade de registros inseridos no SIM de janeiro de 2001 a dezembro de 2011

4.1.2. INCOMPLETUDE

Os resultados da incompletude são apresentados pelas bases SINASC, SIM e SIH através do cálculo abordado na Fórmula 1, e a classificação foi com base na Figura 2.

Foi selecionado o estado de São Paulo para exemplificar os resultados gerados pelas ferramentas automáticas do portal. Demais estados poderão ser vistos e analisados através do site PMAC.

Nestas variáveis verificou-se a porcentagem de não preenchimento. A representação do dado por um traço significa que naquele ano a variável não existia. Valor zero corresponde a uma variável sempre notificada, e valor cem significa que ela nunca foi preenchida.

4.1.2.1. SINASC

Através da análise dos dados, no período em estudo, o Brasil apresentou 32.702.464 registros de nascimento, sendo que em 223.619 foi declarada alguma anomalia congênita.

Nas Tabelas 11 e 12 os dados do estado de São Paulo demonstram que os campos número da DN, código da instalação do Registro, origem, código do município de nascimento, local de nascimento, código do município de residência da mãe, data do nascimento e sexo são sempre preenchidos.

Variáveis classificadas como excelente e relevante no estudo são: idade da mãe, estado civil da mãe, escolaridade da mãe, gestação, gravidez, parto, consultas, hora nascimento, apgar1, apgar5, peso e data cadastro.

Algumas variáveis sofreram variação de classificação ao longo do período de estudo, sendo elas: código de estabelecimento que inicia como péssimo em 2001, em 2003 está como bom e a partir de 2004 é excelente; o código ocupação da mãe é excelente de 2001 a 2005, em 2006 e 2007 é ruim e de 2008 em diante fica regular; quantidade de filho vivo fica entre excelente e bom; quantidade de filho morto varia entre excelente, bom e regular; código bairro residente e bairro nascimento ficam entre ruim e péssimo, com exceção de 2004, que foi excelente; raça cor está de 2001 a 2003 como ruim, de 2004 a 2006 foi regular, 2007 e 2008 bom e a partir de 2009 é excelente.

Na Tabela 12 observa-se que o código anomalia está na classificação péssima. Isto ocorre porque ele não é de preenchimento obrigatório quando há na variável id anomalia os valores indicando ausência de anomalia ou valor ignorado. A inconsistência, tema abordado no próximo item, trará mais informações a respeito deste campo e a relação entre estas variáveis.

Tabela 11 – Porcentagem da incompletude referente a todos os dados disponibilizados do SINASC no período de janeiro 2001 a dezembro 2011

Ano	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
NumeroDN	0	0	0	0	0	0	0	0	0	0	0
Codigo da Instalacao de Registro	-	-	-	-	-	0	0	0	0	0	0
NumeroDV	-	-	-	-	-	-	-	-	-	2,29	0,55
Origem	-	-	-	-	-	-	-	-	-	0	0
PrefixoDN	-	-	-	-	-	-	-	-	-	2,29	0,55
Codigo Cart	-	-	-	-	-	-	-	-	-	94,75	95,92
Numero Registro Cart	-	-	-	-	-	-	-	-	-	99,49	99,72
Data Registro Cart	-	-	-	-	-	-	-	-	-	97,6	98,35
Codigo Estabelecimento	99,86	99,88	14,91	0,17	3,1	3,29	4	0,68	0,63	0,42	0,35
Codigo Municipio Nascimento	0	0	0	0	0	0	0	0	0	0	0
Local Nascimento	0	0	0	0	0	0	0	0	0	0	0
Idade Mae	0,12	0,08	0	0,04	0	0	0	0	0	0	0
Estado Civil Mae	0	0,04	0	0	0,01	2,84	0,24	0,74	0,15	0,3	0,22
Escolaridade Mae	0	0,05	0	0	0,01	3,08	0,58	1,01	0,37	0,39	0,48
Codigo Ocupacao Mae	4,56	2,09	0,92	4,09	3,29	27,38	21,28	18,55	16,39	13,87	13,95
Quantidade Filho Vivo	5,67	0,01	1,07	7,39	8	7,69	5,52	4,15	4,92	2,41	2,5
Quantidade Filho Morto	9,71	0,02	2	15,34	16,74	12,24	10,45	7,93	9,09	4,73	4,49
Codigo Bairro Residente	-	70,59	25,61	0,1	24,05	27,44	31,51	35,11	36,42	-	-
Codigo Bairro Nascimento	-	99,95	23,28	0,05	21,22	99,03	99,1	99,88	99,87	-	-
Codigo Municipio Residente	0	0	0	0	0	0	0	0	0	0	0
Codigo Pais Residente	-	-	-	-	-	-	-	-	-	0,19	0,08
Gestacao	0	0,03	0	0	0	2,68	0,1	0,56	0,16	0,07	0,5
Gravidez	0	0,02	0	0	0	0,06	0,05	0,04	0,04	0,04	0,06
Parto	0	0,02	0	0	0	0,13	0,08	0,08	0,08	0,06	0,08
Consultas	0	0,04	0	0	0,01	0,4	0,44	0,33	0,19	0,08	0,01
Data Nascimento	0	0	0	0	0	0	0	0	0	0	0
Hora Nascimento	-	-	-	-	-	0,39	0,79	0,18	0,11	0,08	0,05
Sexo	0	0	0	0	0	0	0	0	0	0	0
Apgar1	3,54	2,87	0,49	2,13	1,54	4,17	4,68	1,25	0,92	0,55	0,39
Apgar5	3,69	2,79	0,47	2,07	1,47	4,16	4,63	1,21	0,88	0,53	0,36
Raca Cor	24,91	23,87	22,72	19,56	16,39	12,42	8,61	6,65	3,01	1,52	1,43

Tabela 12 – Continuação da Tabela 11 relativa à porcentagem da incompletude, referente a todos os dados disponibilizados do SINASC no período de janeiro 2001 a dezembro 2011

Ano	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Peso	0	0,6	0,62	0,41	0,24	0,72	2,06	0,28	0,26	0,11	0,03
ID anomalia	0,28	0,15	11,44	16,73	10,65	2,63	2,68	1,49	1,05	1	0,94
Data Cadastro	-	-	-	-	-	3,35	3,83	0,49	0,41	0,19	0,08
Codigo anomalia	99,39	99,36	99,35	99,28	99,26	99,3	99,25	99,2	99,06	99,02	98,96
Numero Lote	-	-	-	-	-	-	-	-	-	0,19	0,08
Versao Sistema	-	-	-	-	-	-	-	-	-	0,35	0,08
Data Recebimento	-	-	-	-	-	4,22	6,16	1,45	0,53	0,32	0,16
Dif Data	-	-	-	-	-	-	-	-	-	0,32	0,16
Data recorig	-	-	-	-	-	-	-	-	-	0,48	0,16
Naturalidade Mae	-	-	-	-	-	-	-	-	-	100	83,22
Codigo Municipio Naturalidade	-	-	-	-	-	-	-	-	-	100	83,22
Serie Escolaridade Mae	-	-	-	-	-	-	-	-	-	100	90,21
Data Nascimento Mae	-	-	-	-	-	-	-	-	-	100	82,39
Raca Cor Mae	-	-	-	-	-	-	-	-	-	100	82,46
Quantidade Gestant	-	-	-	-	-	-	-	-	-	100	82,75
Quantidade Parto Normal	-	-	-	-	-	-	-	-	-	100	82,98
Quantidade Parto Cesariana	-	-	-	-	-	-	-	-	-	100	83,04
Idade Pai	-	-	-	-	-	-	-	-	-	100	88,56
Data Ultima Menstruacao	-	-	-	-	-	-	-	-	-	100	86,49
Semana Gestacao	-	-	-	-	-	-	-	-	-	100	82,79
tpmetestim	-	-	-	-	-	-	-	-	-	100	96,78
Consulta Pre-natal	-	-	-	-	-	-	-	-	-	100	82,51
Mes Pre-natal	-	-	-	-	-	-	-	-	-	100	82,93
tpapresent	-	-	-	-	-	-	-	-	-	100	82,51
sttrabpart	-	-	-	-	-	-	-	-	-	100	82,87
stcesparto	-	-	-	-	-	-	-	-	-	100	50,22
tprobson	-	-	-	-	-	-	-	-	-	96,44	76,31
stdnepidem	-	-	-	-	-	-	-	-	-	0,19	0,08
stdnnova	-	-	-	-	-	-	-	-	-	0,19	0
racacor_rn	-	-	-	-	-	-	-	-	-	-	18,97
Codigo Estado - UF	-	0	0	0	0	0	0	0	0	-	-

Além das informações contidas nas Tabelas 11 e 12, tem-se na Figura 8 a análise da incompletude da variável id anomalia em todos os estados ao longo do período estudado, cujos resultados foram divididos por regiões.

A região sul possui a menor média ficando abaixo de 1% com a classificação excelente. A região norte possui a segunda menor média, sendo excelente de 2003 a 2011 nos seus estados. Em 2001 os estados Acre e Amazonas são classificados como bom, e Rondônia é ruim: este em 2002 passa a ser regular. As regiões sudeste e nordeste possuem a média entre excelente e bom, no entanto, destacam-se os estados de São Paulo que, entre 2003 e 2005, foi regular, assim como Minas Gerais em 2001 e esta vai de 2002 a 2005 como bom. O Rio de Janeiro é bom nos anos de 2001 e 2005, a Bahia é regular em 2001 e 2002 e bom entre 2003 a 2010, Sergipe é regular em 2001, 2002 e 2004, sendo bom em 2003. A região centro-oeste é a única média que contém classificação variando entre excelente, bom e regular. Goiás possui dados regulares nos anos de 2003 a 2005. O maior destaque é o Distrito Federal, que é regular em 2001, 2002 e 2007, sendo ruim de 2003 a 2006.

Figura 8 – Representação da incompletude da variável id anomalia no SINASC de janeiro de 2001 a dezembro de 2011

4.1.2.2. SIM

Os dados do SIM, ao longo dos anos de estudo, registraram na base, no Brasil, 11.544.594 registros de óbitos dos quais 141.360 foram declarados com alguma anomalia congênita.

As Tabelas 13 e 14 contêm a totalização do preenchimento das seguintes variáveis: número DO, tipo óbito, data do óbito, sexo, código município residente, local de ocorrência, causa básica, código instalação e origem.

Variáveis classificadas como excelente e relevante no estudo são: data de nascimento, idade, data de atestado, causa básica original, linhaa (todos os campos iniciados com a palavra 'linha' contêm informações complementares do CID), data cadastro e UF (código do estado).

Algumas variáveis que sofreram variações quanto à classificação ao longo do período de estudo são: natural é de 2001 a 2005 excelente, de 2006 a 2009 regular, 2010 ruim e 2011 péssimo; escolaridade ficou em 2001 e 2002 como excelente, de 2003 a 2010 é ruim e em 2011 regular; ocupação de 2001 a 2004 bom, 2005 excelente, 2006 a 2011 ruim; código bairro residente, 2001 péssimo, 2002 a 2005 excelente, 2006 a 2010 ruim; assistência médica, exame, cirurgia e necropsia em 2001 e 2002 excelente, de 2003 a 2011 varia entre ruim e péssimo; linhac vai de 2001 a 2005 como ruim e de 2006 a 2011 regular; linhac varia entre ruim e péssimo; raça cor e estado civil variam entre excelente e bom.

Tabela 13 – Porcentagem da incompletude referente a todos os dados disponibilizados do SIM no período de janeiro 2001 a dezembro 2011

Ano	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
NumeroDO	0	0	0	0	0	0	0	0	0	0	0
Tipo obito	0	0	0	0	0	0	0	0	0	0	0
Data obito	0	0	0	0	0	0	0	0	0	0	0
Hora obito	-	-	-	-	-	6,07	6,67	8,56	5,71	8,64	7,13
Natural	1,22	1,29	1,41	1,21	1,15	15,11	17,26	15,97	16,3	25,64	64,47
Data nascimento	1,13	1,03	0,91	0,76	0,76	0,76	0,73	0,69	0,69	0,63	0,59
Idade	0	0	0	0	0	0,27	0	0	0	0	0
Sexo	0	0	0	0	0	0	0	0	0	0	0
Raca cor	5,19	4,82	1,87	1,36	0,97	5,53	5,24	4,68	4,89	5,21	4,84
Estado civil	0,03	0,03	4,93	0,02	0,1	6,94	6,22	8,14	5,76	5,83	6,51
Escolaridade	0,09	0,09	30,28	33,68	33	31,18	29,4	29,43	28,42	22,34	11,72
Ocupacao	7,2	7,41	8,18	6,93	2	36,25	31,49	27,89	25,79	24,88	24,27
Codigo municipio residente	0	0	0	0	0	0	0	0	0	0	0
Codigo bairro residente	-	100	0,38	0,41	0,5	26,92	33,19	35,18	36,63	38,64	-
Local ocorrencia	0	0	0	0	0	0	0	0	0	0	0
Codigo estabelecimento	99,8	99,79	20,31	19,95	23,41	21,47	20,27	20,6	20,71	20,58	20,59
Codigo municipio ocorrencia	0	0	0	0	0	0	0	0	0	0	0
Codigo bairro ocorrencia	-	-	-	-	-	86,52	87,88	88,99	88,97	88,57	-
Idade mae	97,73	97,99	98	98,08	98,1	97,68	98	97,64	97,66	97,76	97,8
Escolaridade mae	95,58	96	97,86	98,2	98,29	97,78	98,12	97,61	97,64	97,79	97,73
Ocupacao mae	98,01	98,29	98,34	98,39	98,41	98,5	98,6	98,1	98,08	98,16	98,19
Quantidade filho vivo	98,22	98,37	96,14	96,33	96,47	98,03	98,39	97,77	97,75	97,87	97,86
Quantidade filho morto	98,57	98,69	96,15	96,33	96,47	98,34	98,65	97,93	97,87	97,97	97,96
Gravidez	95,58	96	97,36	97,56	97,64	97,42	97,67	97,34	97,43	97,62	97,67
Gestacao	95,58	96	97,38	97,6	97,69	97,44	97,69	97,36	97,44	97,65	97,76
Parto	95,58	96	97,4	97,61	97,69	97,45	97,7	97,36	97,45	97,63	97,68
Obito parto	100	99,99	97,82	97,81	97,87	97,8	97,88	97,48	97,56	97,66	97,74
Peso	97,48	97,69	97,72	97,74	97,76	97,55	97,79	97,45	97,52	97,69	97,76
NumeroDN	-	-	-	-	-	98,34	98,72	98,03	98,04	98,22	98,2
Causa basica original	-	-	-	-	-	1,7	0	0,69	1,08	0,58	0,37
Obito gravidez	93,8	93,88	94,17	97,07	97,09	94,76	94,08	93,76	93,12	92,1	93,33
Obito puerperio	93,8	93,88	97,01	97,21	97,21	94,93	94,28	93,91	93,27	92,23	93,39
Assistencia medica	0,21	0,21	46,75	50,13	52,08	45,06	41,37	41,05	100	100	31,33
Exame	0,13	0,13	31,75	33,67	34,99	31,41	29,26	29,45	31,12	37,09	66,9
Cirurgia	0,14	0,13	31,56	35,18	36,19	32,26	30,09	30,08	31,77	37,59	67,03

Tabela 14 – Continuação da Tabela 13 relativa à porcentagem da incompletude referente a todos os dados disponibilizados do SIM no período de janeiro 2001 a dezembro 2011

Ano	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Necropsia	0,1	0,1	22,82	26,15	27,29	23,64	22,09	21,6	23,66	20,97	23,4
Linhaa	0,01	0,01	0,02	0,01	0,02	1,03	1,35	1,59	1,91	1,36	1,42
Linhad	29,23	28,49	27,16	26,36	24,75	16,64	17,15	17,8	17,73	17,85	18,08
Linhac	52,79	51,94	51,24	50,52	49,99	46,39	46,18	46,38	46,64	47,29	48,44
Data cadastro	-	-	-	-	-	1,9	0,17	1,05	1,63	1,74	1,64
Linhad	81,45	81,48	80,72	80,36	81,01	77,43	76,69	76,42	76,61	77,19	78,28
Linhaii	68,56	66,65	64,72	63,92	61,88	61,34	61,83	61,41	60,55	59,93	59,6
Causa basica	0	0	0	0	0	0	0	0	0	0	0
TP Assina	-	-	-	-	-	100	100	100	-	-	-
Data atestado	-	-	-	-	-	0,2	0,13	0,1	1,1	0,63	0,42
Circunstancia obito	85,76	86,31	87,18	88,16	88,57	90,54	90,59	90,47	90,66	90,89	91,03
Acidente trabalho	94,97	95,1	97,94	95,49	92,08	97,35	97,22	97,03	96,9	96,79	96,79
Fonte	99,89	99,88	91,16	90,39	90,84	91,82	91,95	91,83	91,8	92,02	92
TP pos	-	-	-	-	-	51,28	52,34	48,45	46,74	43,55	43,51
Data investigacao	-	-	-	-	-	94,65	96,14	92,62	91,59	90,66	90,6
Atestante	9,32	9,65	9,71	9,18	23,17	7,42	5,22	4,84	5,24	5,09	7,15
Fonte investigacao	-	-	-	-	-	98,14	94,46	91,86	91,04	89,8	89,97
Data recebimento	-	-	-	-	-	2,73	4,05	100	1,09	0,6	0,36
UF informe	-	0	0	0	0	0	0	0	0	100	-
Codigo instalacao	-	-	-	-	-	0	0	0	0	0	0
CB_pre	-	-	-	-	-	0	0	0	0	0,95	-
Morte parto	-	-	-	-	-	-	-	-	99,74	98,42	98,07
TPobitocor	-	-	-	-	-	-	-	-	97,95	95,91	93,47
Origem	-	-	-	-	-	-	-	-	-	0	0
Data cadastro informacao	-	-	-	-	-	-	-	-	-	98,42	98,07
Data cadastro investigacao	-	-	-	-	-	-	-	-	-	95,91	95,4
NumeroDV	-	-	-	-	-	-	-	-	-	-	0,48
NumeroSUS	-	-	-	-	-	-	-	-	-	-	94,71
Comunsvoim	-	-	-	-	-	-	-	-	-	-	86,18
Data recorig	-	-	-	-	-	-	-	-	-	-	0,36
Data recoriga	-	-	-	-	-	-	-	-	-	-	0,36
Causamat	-	-	-	-	-	-	-	-	-	-	100
Esc2010	-	-	-	-	-	-	-	-	-	-	100
Escmae2010	-	-	-	-	-	-	-	-	-	-	98,81
Stdoepidem	-	-	-	-	-	-	-	-	-	-	0,36
Stdonova	-	-	-	-	-	-	-	-	-	-	0

O SIM contém variáveis com informações dos CIDs, e consequentemente a anomalia congênita. Neste estudo considera-se como principal a variável causa básica, que é classificada como excelente em todos os estados e anos, sendo sempre preenchida.

A segunda variável que se adotou como complementar é a linhaa: em 14 estados é excelente, nos estados do Ceará, Espírito Santo, Goiás, Mato Grosso do Sul, Mato Grosso, Piauí, Rio de Janeiro e Tocantins é excelente e bom. Os estados do Amazonas, Santa Catarina e Sergipe são classificados como excelente, bom e regular. O estado do Rio Grande do Sul é bom e regular. O estado de Roraima tem classificação péssimo em 2003 e nos demais anos é excelente.

4.1.2.3. SIHAC

Para os dados do SIHAC foram selecionados, a partir dos 129.785.443 registros no período, 868.573 registros de internações com anomalias congênitas.

A auditoria das variáveis (60 em 2001, 68 em 2002, 69 em 2003 e 2005, 67 em 2004, 75 em 2006 e 2007, 86 de 2008 a 2010 e 93 em 2011) foi realizada. Uma vez que este banco contém muitos campos, escolheu-se apresentar na Tabela 15 aqueles que tiveram valores de não preenchimento em algum ano ou que foram inseridos ou excluídos ao longo do período de estudo. As demais variáveis que não são apresentadas possuem a totalização do preenchimento de janeiro de 2001 a dezembro de 2011, sendo classificadas como excelente. Dentre elas, as que são relevantes no estudo são: código do município residente, data de nascimento, data de internação, data de saída, sexo, idade, ano de competência, mês de competência, código do estado e morte.

Analisando a tabela também encontra-se na totalização do preenchimento as variáveis: diárias de acompanhante, procedimento solicitado, val_acompanhante, número da rubrica, método contraceptivo 1, método contraceptivo 2, gestante de alto risco, CPF do gestor, código do CNES, nível de complexidade, financiamento, raça cor e diagnóstico principal.

Algumas variáveis que sofreram variações quanto à classificação ao longo do período de estudo são: CGC do hospital é de 2001 a 2007 excelente, de 2008 a 2011 regular; sequencial da AIH foi regular em 2001 e de 2002 a 2011 excelente; vínculo com a previdência iniciou em 2002 como regular, em 2004 foi péssimo, e nos demais anos excelente; código gestor foi de 2002 a 2005 excelente, 2006 regular, 2007 bom, 2008 a 2011 péssimo; status de infecção hospitalar, CID asso e CID morte iniciaram 2006 a 2007 excelente e de 2008 a 2011 ficaram péssimas; diagnóstico secundário é péssimo todos os anos; CNPJ da mantenedora inicia em 2008 e até 2011 é ruim; as variáveis número do procedimento, CPF do auditor, grau de instrução, CID notificação, código brasileiro de ocupação e código de acidente de trabalho variam, em alguns anos são excelente, em outros péssimo.

Tabela 15 – Porcentagem da incompletude referente às variáveis disponibilizadas pelo SIHAC no período de janeiro 2001 a dezembro 2011.

Ano	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
cgc do hospital	0	0	0	0	0	0	0	16,24	14,95	13,75	14,21
diarias de acompanhante	-	-	-	-	0	0	0	0	0	0	0
procedimento solicitado	-	-	-	-	-	0	0	0	0	0	0
val_acompanhante	-	-	-	-	0	0	0	0	0	0	0
diagnóstico principal	0	0	0	0	0 0,01	-	0	0	0	0	0
diagnóstico secundário	84,19	79,49	76,87	77,41	75,99	76,08	75,02	80,25	79,21	77,39	76,99
numero da rubrica	-	-	-	-	-	0	0	0	0	0	0
ind_vdrl	-	-	-	-	-	9,46	0	0	0	0	0
numero do procedimento	0	0	0	0	0	88,9	100	100	100	100	100
cpf do auditor	0	0	0	0	0	0	0	100	100	100	100
grau de instrução	0	0	0	0	0 0,01	-	0	0	0	0	0
cid notificação	80,09	0	0	0	0 0,01	-	0	100	100	100	100
método contraceptivo 1	0	0	0	0	0 0,01	-	0	0	0	0	0
método contraceptivo 2	0	0	0	0	0 0,01	-	0	0	0	0	0
gestante de alto risco	0	0	0	0	0 0,01	-	0	0	0	0	0
sequencial da aih	17,48	0	0	0	0	0	0	0	0	0	0
código brasileiro de ocupação	-	0	0	62,34	0	0	0	0	0	0	0
código de acidente de trabalho	-	0	0	62,34	0	0	0	0	0	0	0
vinculo com a previdencia	-	11,1	0	62,34	0	0	0	0	0	0	0
código do gestor	-	0	0	0	0	11,36	7,06	100	100	100	100
gestor_tp	-	-	-	-	-	-	-	0	0	0	0
cpf do gestor	-	0	3,96	0	0	0	0	0	0	0	0
data da autorização dada pelo gestor	-	0	27,79	0	0	87,25	98,39	100	100	100	100
código do cnes	-	-	0	0	0	0	0	0	0	0	0
cnpj da mantenedora	-	-	-	-	-	-	-	39,59	38,14	35,93	34,64
status de infecção hospitalar	-	-	-	-	-	0	0	100	100	100	100
cid asso	-	-	-	-	-	0	0	96,17	95,06	94,26	94,75
cid morte	-	-	-	-	-	0	0	96,11	95,95	96,28	96,41
nível de complexidade	-	-	-	-	-	-	-	0	0	0	0
financiamento	-	-	-	-	-	-	-	0	0	0	0
faec_tp	-	-	-	-	-	-	-	99,54	99,76	99,78	99,71
regct	-	-	-	-	-	-	-	0	0	0	0
raça_cor	-	-	-	-	-	-	-	0	0	0	0
etnia	-	-	-	-	-	-	-	100	100	68,27	0

O SIHAC contém variáveis com informações dos CIDs e consequentemente a anomalia congênita. Neste estudo considera-se como principal a variável diagnóstico principal que é classificada como excelente em todos os estados e anos. A segunda variável complementar é o diagnóstico secundário que, embora seu preenchimento seja péssimo, ele apenas auxiliou agregando mais notificações quando a informação de AC existia no registro da AIH.

4.1.3. INCONSISTÊNCIA

A inconsistência gerada a partir da Fórmula 2 e apresentada na Figura 9 contém as variáveis referente à AC e, assim como a incompletude, é adotada a classificação com base na Figura 2.

Os resultados demonstram que no período de 2001 a 2005 a média é constante com valores próximos a nulo, ou seja, todas as regiões se enquadram na classificação “excelente”. A partir do ano de 2006 notam-se picos de inconsistência, cujos mais expressivos são: no ano de 2007, os estados de Roraima e Paraíba nas regiões norte e nordeste; em 2008, os estados

do Espírito Santo e Paraná nas regiões sudeste e sul; no ano de 2009, Mato Grosso na região centro-oeste.

Analisando o maior resultado ocorrido em 2007 na Paraíba com valor 20 e lembrando que a base adotada no cálculo corresponde a 10.000, tem-se a classificação “excelente” neste caso e em todos os demais, ou seja, a inconsistência entre as variáveis é pequena e próxima a zero.

Figura 9 – Inconsistência nos campos com notificação de AC na base SINASC de janeiro de 2001 a dezembro de 2011.

4.2. COEFICIENTES EM EPIDEMIOLOGIA

Os resultados dos coeficientes e taxas são apresentados nesta ordem: coeficiente de incidência, coeficiente de mortalidade infantil, taxa de mortalidade perinatal, taxa de mortalidade neonatal tardia e mortalidade segundo causas. Todos são ilustrados por gráficos divididos pelas regiões norte, nordeste, centro-oeste, sudeste e sul.

4.2.1. COEFICIENTE DE INCIDÊNCIA

Os resultados do coeficiente de incidência de notificação de AC demonstram um crescimento ao longo dos anos de estudo. A média das regiões norte, nordeste, centro-oeste, sudeste e sul iniciam respectivamente com 29, 37, 42, 53 e 68 em 2001 e finalizam com 50, 69, 66, 79 e 88 em 2011.

A base adotada no cálculo deste coeficiente corresponde a 10.000. Ao selecionar o maior dado ocorrido em 2011 no estado de São Paulo com valor 104, tem-se uma incidência baixa, ou seja, um caso de incidência de notificação de AC a cada cem registros preenchidos.

A média das médias em 2001 equivale a 45,8 e ao final de 2011 corresponde a 70,4. Estes dados demonstram que em 2001, no Brasil, a média de incidência de notificação estava em 0,458% e, em 2011, era 0,704%, um crescimento de 54% ao longo dos anos de estudo.

Figura 10 – Coeficiente de Incidência de notificação de AC na base do SINASC de janeiro de 2001 a dezembro de 2011.

4.2.2. COEFICIENTE DE MORTALIDADE INFANTIL

Os resultados representados na Figura 11 permite observar uma diminuição do coeficiente de mortalidade infantil ao longo dos anos estudados. A média das regiões norte, nordeste, centro-oeste, sudeste e sul inicia respectivamente em 2001 com valores representados por 22, 24, 18, 17 e 16 e em 2011 finalizam com 15, 15, 13, 12 e 11. A região sul conta com os menores valores de coeficiente de mortalidade infantil desde o início do estudo até o final do período. Em contraposição, a região nordeste, que em 2001 tinha os maiores valores e se distanciava da região sul, finaliza em 2011 com uma média próxima das demais regiões. Uma das peculiaridades do CMI ocorreu no estado do Rio de Janeiro, que apresentou um pico em 2006.

A média das médias em 2001 equivale a 19,4 e ao final de 2011 corresponde a 13,2. Estes dados demonstram que no Brasil houve uma diminuição de 31,9% ao longo dos anos de estudo.

Figura 11 – Coeficiente de Mortalidade Infantil na base do SIM de janeiro de 2001 a dezembro de 2011.

4.2.3. TAXA DE MORTALIDADE PERINATAL

Os resultados contidos na Figura 12 permitem observar uma diminuição da taxa de mortalidade perinatal nos estados ao longo do período de estudo. A média das regiões norte, nordeste, centro-oeste, sudeste e sul inicia em 2001 com valores 23, 24, 21, 19 e 16 e em 2011 finalizam com 16, 18, 15, 15 e 13. A TMP apresenta características semelhantes ao CMI, com a região sul permanecendo com os menores valores, e no sudeste, em 2006, com um pico

acentuado no Rio de Janeiro. A média das médias em 2001 foi ‘20,6’ e em 2011 foi ‘15,4’, ou seja, no Brasil a TMP iniciou com cerca de vinte a cada mil registros de nascimento e óbitos fetais e no final do período estava com quinze para cada mil registros. Esta diminuição representa cerca de 25% ao longo do período de estudo.

Figura 12 – Taxa de Mortalidade Perinatal na base do SIM de janeiro de 2001 a dezembro de 2011.

4.2.4. TAXA DE MORTALIDADE NEONATAL TARDIA

A Figura 13 apresenta os resultados da taxa de mortalidade neonatal tardia e inicialmente observa-se certa constância nos dados ao longo dos anos de estudo. A média das regiões norte, nordeste, centro-oeste, sudeste e sul varia pouco e permanece entre os valores dois e três para cada mil registros de nascimento.

Figura 13 – Taxa de Mortalidade Neonatal Tardia na base do SIM de janeiro de 2001 a dezembro de 2011.

4.2.5. MORTALIDADE SEGUNDO CAUSAS

A representação dos resultados da mortalidade segundo causas de AC está ilustrada na Figura 14. Nestes dados observa-se certa constância da taxa ao longo dos estados no período de estudo. A média das regiões norte, nordeste, centro-oeste, sudeste e sul inicia respectivamente em 2001 com ‘2,3’, ‘1,7’, ‘2,9’, ‘2,6’ e ‘2,9’ e em 2011 finalizam com ‘2,6’, ‘2,8’, ‘2,8’, ‘2,7’ e ‘2,8’. A base adotada neste cálculo corresponde a 1.000, ou seja, em média no Brasil tem-se que, de mil registros de mortalidade, aproximadamente dois contêm informações de AC.

Figura 14 – Mortalidade Segundo Causas de AC na base do SIM de janeiro de 2001 a dezembro de 2011.

4.3. RELACIONAMENTO PROBABILÍSTICO

Como visto na seção 3.7, no pareamento foram utilizadas as variáveis que contêm o CID do capítulo XVII. Desta forma todos os registros que não contêm esta informação foram descartados. Esta decisão ocorreu para reduzir ainda mais o processamento e melhorar no desempenho do software. Nesta técnica do relacionamento probabilístico foi gerado arquivos contendo somente informações de AC, ou seja, além dos SIHAC, teve-se a redução do SINASC para SINASCAC e do SIM para SIMAC, cobrindo todo o território nacional no período de janeiro 2001 a dezembro 2011. Esta redução foi realizada apenas nesta técnica e não afetou as bases principais, pois uma nova base foi gerada com os resultados do linkage.

Após as etapas de padronização e relacionamento (definidos a blocagem e pareamento) o software OpenRecLink gerou para cada estado um arquivo com os seguintes cálculos: probabilidade m_i do campo concordar dado que é um par verdadeiro; probabilidade u_i do campo concordar dado que é um par falso; probabilidade p de ser par; frequência de concordância; pesos da concordância; pesos da discordância; poder discriminante do campo; padrões, frequências e scores; estes valores estão ilustrados na Figura 15. Em destaque estão os limiares inferiores e superiores que são usados na última etapa.

Após gerar os cálculos e executar o relacionamento, entrou-se na etapa final denominada combina, que está ilustrada na Figura 16. Destacaram-se nesta imagem algumas partes e elas foram enumeradas e descritas a seguir:

- **Item 1:** representa a chave de blocagem (sexo, ano de nascimento e código município residente), seguido dos campos de pareamento (data de nascimento e CID do capítulo XVII). Neste caso usou-se uma ilustração do relacionamento SIM-SINASC. Lembrando que as chaves de blocagem sempre são iguais entre as bases, o que pode variar são os caracteres das variáveis do pareamento e esta diferença altera o valor do score;
- **Item 2:** precisa ser preenchido pelos valores do limiar superior e inferior e neste estudo usaram-se os resultados gerados pelos arquivos de saída e que estão em destaque na Figura 15;
- **Item 3:** mostra duas colunas, na primeira há o valor do score e na segunda a quantidade de arquivos com aquele score, por exemplo: para o score 15,926597554743 houve 122 arquivos relacionados.

Ao final da etapa combina, há os arquivos relacionados com seus respectivos scores.

Parâmetros de saída

Número de iterações do algoritmo EM: 15

Menor diferença da log-verossimilhança entre iterações: 0,00094422

Maior valor da função de log-verossimilhança: -32,0988

Parâmetros estimados

Probabilidade $m(i)$ do campo concordar dado que é um par:

$m(1) = 0,999877$

$m(2) = 0,46526$

$m(3) = 0,960868$

Probabilidade $u(i)$ do campo concordar dado que é um não-par:

$m(1) = 0,119891$

$m(2) = 0,00153596$

$m(3) = 0,00801207$

Probabilidade (p) de ser par: 0,00477782

Frequência de concordância (f) do campo i :

$f(1) = 1231879$

$f(2) = 15782$

$f(3) = 82324$

Limiar inferior (na escala do escore): -2,50474

Limiar superior (na escala do escore): 9,06518

Pesos da concordância $A(i)$ $[\log(w_i/u_i)]$:

$A(1) = 3,06003$

$A(2) = 8,24275$

$A(3) = 6,90602$

Pesos da discordância $D(i)$ $[\log((1-m_i)/(-u_i))]$:

$D(1) = -12,8106$

$D(2) = -0,900874$

$D(3) = -4,66389$

Poder discriminante do campo $P(i)$ $[A(i)-D(i)]$:

$P(1) = 15,8706$

$P(2) = 9,14363$

$P(3) = 11,5699$

Padrões, frequências e escores observados:

0 / 0 / 0 / score: -18,3753

0 / 0 / 1 / score: -6,80543

0 / 0 / 0 / score: -18,3753

0 / 0 / 0 / score: -18,3753

1 / 0 / 0 / score: -2,50474

1 / 0 / 1 / score: 9,06518

1 / 1 / 0 / score: 6,63889

1 / 1 / 1 / score: 18,2088

Figura 15 – Exemplo de arquivo gerado pelo software openRecLink com os cálculos da etapa do relacionamento

Referência

	Valor
NUMERODN	20139229
DTNASC	25082006
SEXO	1
ANONASC	2006
CODMUNRES	421220
CODANOMAL	Q000
IDANOMAL	1
LOCNASC	1
CODBAIRES	00000002
CODESTAB	2270162

Comparação

	Valor
NUMERODO	03272369
DTNASC	25082006
SEXO	1
ANONASC	2006
CODMUNRES	421220
LINHAA	Q000
CAUSABAS	Q000
TIPOBITO	2
DTOBITO	27082006
UNDAOBITO	0050

Score do par: 15,926597554743

Status:

- ☒ Par
- ☐ Dúvida
- ☐ Não par
- ☐ Ignorado

Score + min: 0

Score - máx: 0

Marca todos

Limpa todos

	Freq.
15,926597554743	122
15,557953722351	15
15,189309889959	29
9,239123952181	22
9,111168444109	13
8,870480110780	

Fim

Avança

Volta

Início

Passo: 0

✓ ✗

Figura 16 – Ilustração da etapa combinada gerada pelo software openRecLink

O produto gerado ao final de todo o processo da técnica de relacionamento probabilístico foram arquivos por estado com a junção dos campos das duas bases que posteriormente foram inseridos e viraram uma nova base de dados.

Na Figura 17 são apresentadas as informações quantitativas do relacionamento SIHAC-SINASCAC, divididas por regiões do país. Lembrando que apenas uma DN pode ser gerada por pessoa, no entanto um indivíduo pode sofrer mais de uma internação. Estas observações justificam alguns conjuntos de linkage nos quais foram encontrados um registro do SINASC pareando com diferentes registros do SIH. A maioria dos estados apresentou valores do relacionamento abaixo da quantidade de registros das bases originais. Houve duas exceções: o Distrito Federal e São Paulo, dos quais os valores do relacionamento ultrapassaram os valores das bases.

Figura 17 – Relacionamento probabilístico SIHAC-SINASCAC

Na Figura 18, encontram-se as informações quantitativas do relacionamento SIMAC-SINASCAC, também divididas por regiões do país. Neste relacionamento tem-se a mesma característica da DN, ou seja, um indivíduo só pode ter uma DO. No geral, a quantidade de registros gerados nos resultados do linkage ficou abaixo das quantificações das bases relacionadas.

Figura 18 – Relacionamento probabilístico SIMAC-SINASCAC

Na sequência, há análises de rastreabilidade na tentativa de entender o fluxo (comportamento migratório) dos registros com notificação de AC em todos os estados. Os resultados referentes à rastreabilidade são apresentados em forma de porcentagem.

Na Tabela 16, apresentam-se os resultados dos três bancos públicos. Na segunda coluna, há os dados do SIHAC com a frequência de registros de indivíduos que residem e foram internados na mesma cidade. Na terceira coluna, há aqueles que residem numa cidade, mas foram internados em outra. Na quarta coluna, há os dados do SINASCAC que indicam registros de nascimentos de indivíduos que residem e nasceram numa mesma cidade. Na quinta coluna, há informações diferentes quanto à cidade de residência e nascimento. Na sexta coluna, há os dados do SIMAC com registros de indivíduos que nasceram e o desfecho foi na mesma cidade. Na sétima coluna, há informações de registros com nascimento e desfecho em

cidades diferentes. Nas Tabelas 16 e 17, o critério dos tons de cinza na coluna estados foi a divisão por região do país:

Tabela 16 – Resultados da rastreabilidade dos dados do SIHAC, SINASCAC e SIMAC.

Estados	Cidade(%) SIHAC	Migração(%) SIHAC	Cidade(%) SINASCAC	Migração(%) SINASCAC	Cidade(%) SIMAC	Migração(%) SIMAC
Acre	72,99908004	27,00091996	79,1666667	20,83333333	62,78317152	37,21682848
Amapá	73,77300613	26,22699387	81,4696486	18,53035144	64,04293381	35,95706619
Amazonas	93,31704337	6,682956628	94,8223156	5,177684371	81,46385004	18,53614996
Pará	63,80110606	36,19889394	75,1698015	24,83019854	50,11240165	49,88759835
Rondônia	62,8919218	37,1080782	78,0088217	21,99117832	55,83266291	44,16733709
Roraima	70,87452471	29,12547529	75,6190476	24,38095238	64,31623932	35,68376068
Tocantins	50,91477019	49,08522981	58,672799	41,32720105	42,27799228	57,72200772
NORTE (média)	69,79592176	30,20407824	77,5613001	22,43869991	60,1184645	39,8815355
Alagoas	45,34617897	54,65382103	56,7361801	43,26381993	45,88433938	54,11566062
Bahia	63,65336974	36,34663026	70,9248596	29,0751404	53,12052834	46,87947166
Ceará	52,69056963	47,30943037	71,3062852	28,69371482	54,54688236	45,45311764
Maranhão	62,57049402	37,42950598	74,2597403	25,74025974	47,4680549	52,5319451
Paraíba	44,98611792	55,01388208	47,2474868	52,75251316	42,68939394	57,31060606
Pernambuco	29,17218017	70,82781983	49,2697176	50,73028238	36,61934101	63,38065899
Piauí	40,52478134	59,47521866	69,4840834	30,51591658	55,6474969	44,3525031
Rio Grande do Norte	42,28618572	57,71381428	58,8790561	41,12094395	49,28087986	50,71912014
Sergipe	35,01390176	64,98609824	43,3407738	56,65922619	38,28996283	61,71003717
NORDESTE (média)	46,24930881	53,75069119	60,1609092	39,8390908	47,06076439	52,93923561
Distrito Federal	57,00167906	42,99832094	98,3755195	1,624480544	96,26485569	3,735144312
Goiás	47,29778581	52,70221419	69,4026062	30,59739381	38,79781421	61,20218579
Mato Grosso	59,85899214	40,14100786	70,458221	29,54177898	44,86166008	55,13833992
Mato Grosso do Sul	68,32019406	31,67980594	86,1378205	13,86217949	59,63718821	40,36281179
CENTRO-OESTE (média)	58,11966278	41,88033723	81,0935418	18,90645821	59,89037955	40,10962045
Espírito Santo	30,19721234	69,80278766	56,812933	43,18706697	41,10918544	58,89081456
Minas Gerais	52,31395039	47,68604961	65,6714223	34,32857773	51,10740653	48,89259347
Rio de Janeiro	63,92751012	36,07248988	75,383395	24,61660497	64,73601945	35,26398055
São Paulo	51,04471367	48,95528633	75,2494379	24,75056211	63,97051724	36,02948276
SUDESTE (média)	49,37084663	50,62915338	68,2792971	31,72070295	55,23078218	44,76921783
Paraná	44,0476404	55,9523596	72,9699309	27,03006914	53,26347683	46,73652317
Rio Grande do Sul	43,32335099	56,67664901	70,8392031	29,16079694	51,94022656	48,05977344
Santa Catarina	42,98859173	57,01140827	66,1781145	33,82188548	48,10888942	51,89111058
SUL (média)	43,45319437	56,54680563	69,9957495	30,00425052	51,1041976	48,8958024

Na Tabela 17, apresentam-se os resultados dos dois relacionamentos feitos neste projeto, sendo eles o SIHAC-SINASCAC e SIMAC-SINASCAC. As variáveis usadas como critério nestes resultados foram: código do município de nascimento para o SINASCAC, código do município de movimentação para o SIHAC e município de ocorrência para o SIMAC. Ao compará-las, quando iguais correspondem à mesma cidade, quando diferem correspondem à migração. Lembrando que na blocagem usamos o código de município residente e este sempre será igual e, por isto, não foi utilizado nesta análise.

Tabela 17 – Continuação dos resultados da rastreabilidade dos dados dos linkage SIHAC-SINASCAC e SIMAC-SINASCAC.

Estados	Cidade(%)	Migração(%)	Cidade(%)	Migração(%)
	SIHAC-SINASCAC	SIHAC-SINASCAC	SIMAC-SINASCAC	SIMAC-SINASCAC
Acre	86,59793814	13,40206186	100	0
Amapá	64,28571429	35,71428571	65,38461538	34,61538462
Amazonas	98,23529412	1,764705882	97,67950963	2,320490368
Pará	81,16800677	18,83199323	96,85534591	3,144654088
Rondônia	67,79661017	32,20338983	94	6
Roraima	91,30434783	8,695652174	88,73239437	11,26760563
Tocantins	67,2489083	32,7510917	87,5	12,5
NORTE (média)	79,51954566	20,48045434	90,02169504	9,978304959
Alagoas	84,75452196	15,24547804	89,88505747	10,11494253
Bahia	92,1664218	7,833578202	93,57309357	6,426906427
Ceará	76,34943182	23,65056818	94,88536155	5,114638448
Maranhão	82,60869565	17,39130435	91,57088123	8,429118774
Paraíba	87,48068006	12,51931994	98,29059829	1,709401709
Pernambuco	80,75588122	19,24411878	85,62691131	14,37308869
Piauí	68,09116809	31,90883191	93,24894515	6,751054852
Rio Grande do Norte	77,17842324	22,82157676	96,58385093	3,416149068
Sergipe	77,50556793	22,49443207	97,4137931	2,586206897
NORDESTE (média)	80,76564353	19,23435647	93,45316584	6,546834156
Distrito Federal	82,30080806	17,69919194	91,67180753	8,328192474
Goiás	66,24902114	33,75097886	94,11764706	5,882352941
Mato Grosso	71,71903882	28,28096118	81,96721311	18,03278689
Mato Grosso do Sul	74,46808511	25,53191489	91,02564103	8,974358974
CENTRO-OESTE (média)	73,68423828	26,31576173	89,69557718	10,30442282
Espírito Santo	43,75876578	56,24123422	58,60683188	41,39316812
Minas Gerais	71,62434502	28,37565498	86,4978903	13,5021097
Rio de Janeiro	90,15840042	9,841599585	90,58450907	9,41549093
São Paulo	85,83939558	14,16060442	91,05259682	8,947403184
SUDESTE (média)	72,8452267	27,1547733	81,68545703	18,31454298
Paraná	63,53144836	36,46855164	90,32983508	9,670164918
Rio Grande do Sul	72,65126178	27,34873822	88,6374502	11,3625498
Santa Catarina	51,46797153	48,53202847	80,21978022	19,78021978
SUL (média)	62,55022723	37,44977277	86,3956885	13,6043115

Observa-se que a menor média de migração corresponde aos nascimentos: esta característica se dá em todas as regiões. A região norte possui as menores médias comparada às demais regiões, seguida por centro-oeste. As regiões nordeste, sudeste e sul ficam com as médias muito próximas umas das outras. Nos linkages percebe-se a média baixa para as migrações. Isto indica que, ao relacionar o conjunto de campos na blocagem e pareamento, aqueles que eram da mesma cidade e nela ficaram para internação ou desfecho do óbito foram mais bem relacionados.

4.4. PORTAL DE MONITORAMENTO DE ANOMALIA CONGÊNITA

O resultado do desenvolvimento e implantação do sistema para a gestão da informação e vigilância de AC no país é mostrado a seguir através das interfaces gráficas. Ao final deste capítulo temos a apresentação dos testes de avaliação.

4.4.1. APRESENTAÇÃO DO PMAC

O link de acesso ao site corresponde a: <http://ciis.fmrp.usp.br/anomalia>. Ao abrir o portal, o primeiro módulo denominado menu “Início” é apresentado e traz informações básicas para auxiliar na navegação do usuário, como ilustrado na Figura 19.

Figura 19 – Interface do menu início do PMAC

4.4.2. MÓDULO HISTÓRICO

O segundo módulo corresponde ao Histórico, que tem por finalidade situar, de forma geral e simples, todos os aspectos e informações relevantes para auxiliar o usuário quando este selecionar as ferramentas. Na sequência de imagens, há as Figuras 20, 21, 22, 23 e 24, que representam respectivamente o Histórico de Informações Gerais, Histórico SIHAC, Histórico SINASC, Histórico SIM e Histórico do Linkage.

Figura 20 – Interface do submenu Histórico de Informações Gerais.

Figura 21 – Interface do submenu Histórico SIHAC.

Figura 22 – Interface do submenu Histórico SINASC.

Figura 23 – Interface do submenu Histórico SIM.

Figura 24 – Interface do submenu Histórico do Linkage.

4.4.3. FERRAMENTAS SEMELHANTES ENTRE SINASC, SIHAC E SIM

O terceiro, quarto e quinto módulos, correspondentes à Ferramenta SIHAC, Ferramenta SINASC e Ferramenta SIM, possuem em comum três submenus apresentados a seguir:

- a) **Busca de Dados:** permite pesquisar informações das variáveis de cada sistema através de seleção da consulta por nível estadual ou municipal, sexo, período desejado e o CID da AC (por Grupo, Categoria ou Subcategoria). Ver exemplo nas Figuras 25 e 26;

- b) **Auditoria Automática:** permite auditar as variáveis de cada sistema, através do cálculo da incompletude que corresponde à taxa de campos (valores) em branco de um registro. Pode-se, para isto, selecionar por nível estadual ou municipal, o período desejado e as variáveis a serem auditadas. Ver exemplo nas Figuras 27 e 28;
- c) **Frequência:** permite verificar a quantidade de registros e sua frequência com AC. Pode-se, para isto, consultar por nível estadual ou municipal, sexo, período desejado e o CID da AC (por Grupo, Categoria ou Subcategoria). Ver exemplo nas Figuras 29 e 30.

Busca de Dados - SIHAC

A ferramenta provê o Perfil Nosológico de pacientes com Anomalia Congênita tendo com base na Autorização de Internação Hospitalar (AIH). Assim, é permitido abaixo selecionar sua consulta:

- (a) em nível estadual;
- (b) em nível municipal;
- (c) por sexo;
- (d) em todo o período estudado ou em um ano específico e
- (e) envolvendo todas as anomalias congênitas por grupo, categoria ou uma anomalia específica.

Localidade:

Estado:

Município:

Sexo:
☐ Feminino ☐ Masculino ☒ Ambos

Ano de Nascimento:

CID - Anomalia Congênita:

Grupo:

Categoria:

Subcategoria:

Figura 25 – Interface do submenu Busca de Dados no banco SIHAC.

Resultado - Busca de Dados - SIHAC											
<div>Voltar</div> <div> - PESQUISA REALIZADA - Estado: SÃO PAULO Município de Referência: RIBEIRÃO PRETO Ano de Internação: 2010 Sexo: AMBOS CID: TODOS ! </div>											
NumeroAIH	MunicRes	MunicMov	DataNasc	Sexo	DataInter	DataSaida	DiagPrinc	DiagSecun	Idade	Morte	Raca
3510102350744	RIBEIRÃO PRETO	RIBEIRÃO PRETO	20100115	Masculino	20100129	20100202	Q038	Q048	14	0	01
3510102343682	RIBEIRÃO PRETO	RIBEIRÃO PRETO	20100101	Feminino	20100102	20100108	Q213		1	0	02
3510102363911	RIBEIRÃO PRETO	RIBEIRÃO PRETO	20100108	Masculino	20100302	20100302	Q668		1	0	01
3510102364318	RIBEIRÃO PRETO	RIBEIRÃO PRETO	20100119	Feminino	20100202	20100208	Q423	P599	14	0	01
3510102364307	RIBEIRÃO PRETO	RIBEIRÃO PRETO	20100119	Feminino	20100120	20100202	Q212	P221	1	0	01
3510104486944	RIBEIRÃO PRETO	RIBEIRÃO PRETO	20100210	Feminino	20100311	20100323	Q250	Z380	1	0	02
3510104487912	RIBEIRÃO PRETO	RIBEIRÃO PRETO	20100323	Feminino	20100323	20100325	Q633	Z380	0	0	01
3510104488187	RIBEIRÃO PRETO	RIBEIRÃO PRETO	20100312	Masculino	20100312	20100322	Q379	Z380	0	0	01
3510104493005	RIBEIRÃO PRETO	RIBEIRÃO PRETO	20100223	Feminino	20100223	20100225	Q270	Z380	0	0	01
3510104496635	RIBEIRÃO PRETO	RIBEIRÃO PRETO	20100223	Masculino	20100304	20100309	Q211	P211	9	0	01
3510108919779	RIBEIRÃO PRETO	RIBEIRÃO PRETO	20100302	Feminino	20100302	20100505	Q221	Q250	0	0	01
3510105235692	RIBEIRÃO PRETO	RIBEIRÃO PRETO	20100319	Masculino	20100323	20100327	Q532		4	0	01
3510108909384	RIBEIRÃO PRETO	RIBEIRÃO PRETO	20100409	Masculino	20100412	20100416	Q620		3	0	01
3510109886349	RIBEIRÃO PRETO	RIBEIRÃO PRETO	20100327	Masculino	20100414	20100420	Q250	P073	18	0	01
3510108927633	RIBEIRÃO PRETO	RIBEIRÃO PRETO	20100411	Feminino	20100413	20100416	Q211	Q250	2	0	01
3510109882610	RIBEIRÃO PRETO	RIBEIRÃO PRETO	20100302	Feminino	20100302	20100305	Q012	G919	0	0	01
3510108921132	RIBEIRÃO PRETO	RIBEIRÃO PRETO	20100410	Masculino	20100410	20100421	Q210	Z380	0	0	03

Figura 26 – Interface dos resultados do submenu Busca de Dados no banco SIHAC.

Auditoria Automática- SINASC	
<p>A ferramenta provê uma verificação de incompletude que corresponde à taxa de campos (valores) em branco de um registro, ou seja, uma informação incompleta ou não notificada de cada variável. O resultado segue a lógica da fórmula apresentada a seguir:</p> <p>(*) Incompletude = {(Quantidade de registros em branco) / (Total de registros)}.</p> <p>Assim, é permitido abaixo selecionar sua consulta:</p> <p>(a) em nível estadual;</p> <p>(b) em nível municipal;</p> <p>(c) em todo o período estudado ou em um ano específico;</p> <p>(d) por variável ou um conjunto delas disponibilizadas no banco de dados;</p>	
<p>Localidade:</p> <p>Estado: <input type="text" value="SÃO PAULO"/></p> <p>Município: <input type="text" value="RIBEIRÃO PRETO"/></p> <p>Ano de Nascimento: <input type="text" value="Todos os Anos"/></p> <p>Variáveis:</p> <p>☐ Escolher ☒ Todas</p> <p>Campos:</p> <p><input type="text" value="NumeroDN"/></p> <p><input type="text" value="Codigo da Instalacao de Registr"/></p> <p><input type="text" value="NumeroDV"/></p> <p><input type="text" value="Origem"/></p> <p><input type="button" value="Buscar"/></p>	
<p>Obs: Uma análise foi realizada pela equipe do PMAC e pode ser visualizada clicando a seguir: <input type="button" value="Visualizar"/></p>	

Figura 27 – Interface do submenu Auditoria Automática no banco SINASC.

Análise - Auditoria de Não Preenchimento dos dados no Sistema de Nascidos Vivos.

Voltar

- PESQUISA REALIZADA -
Estado: SÃO PAULO
Município de Nascimento:
RIBEIRÃO PRETO
Ano de Nascimento: 2001 a 2011

Ano	Total	numerodn	codinst	numerodv	origem	prefixodn	codcart	numregcart	dtregcart	codestab	codmunasc	locnasc	idadema	estcivma
2001	9660	0	-	-	-	-	-	-	-	100	0	0	0.1	
2002	9833	0	-	-	-	-	-	-	-	100	0	0	0.02	0.0
2003	9813	0	-	-	-	-	-	-	-	15.65	0	0	0	
2004	9775	0	-	-	-	-	-	-	-	0.19	0	0	0.01	
2005	10203	0	-	-	-	-	-	-	-	0.41	0	0	0	
2006	9995	0	0	-	-	-	-	-	-	0.25	0	0	0	0.0
2007	9925	0	0	-	-	-	-	-	-	0.32	0	0	0	0.0
2008	10448	0	0	-	-	-	-	-	-	0.21	0	0	0	0.0
2009	10701	0	0	-	-	-	-	-	-	0.33	0	0	0	0.1
2010	11247	0	0	0.03	0	0.03	100	99.96	2	0.17	0	0	0	0.0
2011	11728	0	0	0.03	0	0.03	99.98	99.98	33.36	0.22	0	0	0	0.0

Download dos dados

Obs: Os valores são apresentados em frequência de registros que não contém preenchimento, o campo denominado Total corresponde ao valor absoluto representando 100%.

Figura 28 – Interface dos resultados do submenu Auditoria Automática no banco SINASC.

Frequência de Anomalias Congênitas - SIM

A ferramenta provê os dados com a frequência de pacientes com Anomalia Congênita tendo como base a Declaração de Óbito (DO). Assim, é permitido abaixo selecionar sua consulta:

- (a) em nível estadual;
- (b) em nível municipal;
- (c) por sexo;
- (d) em todo o período estudado ou em um ano específico e
- (e) envolvendo todas as anomalias congênitas por grupo, categoria ou uma anomalia específica.

Localidade:

Estado:
SÃO PAULO ▼

Município:
RIBEIRÃO PRETO ▼

Sexo:
☐ Feminino
 ☐ Masculino
 ☒ Ambos

Ano de Nascimento:
 2008 ▲
 2009
 2010 ▼
 2011

CID - Anomalia Congênita:

Grupo:
Todos os CIDs ▼

Categoria:
-- Escolha um Grupo -- ▼

Subcategoria:
-- Escolha uma Categoria -- ▼

Buscar

Figura 29 – Interface do submenu Frequência no banco SIM.

Resultado - Frequência de Anomalias Congênitas - SIM				
<div>Voltar</div> <div> - PESQUISA REALIZADA - Estado: SÃO PAULO Município de Referência: RIBEIRÃO PRETO Ano do Óbito: 2010 Sexo: AMBOS CID: TODOS ! </div>				
CIDAnomalia	NomeAnomalia	Quantidade	FrequenciaEntreAC	FrequenciaGeral
Q913	Síndrome de Edwards não especificada	8	8.88889	0.1483679525
Q203	Comunicação ventrículo-atrial discordante	6	6.66667	0.1112759644
Q249	Malformação não especificada do coração	6	6.66667	0.1112759644
Q213	Tetralogia de Fallot	5	5.55556	0.0927299703
Q909	Síndrome de Down não especificada	4	4.44444	0.0741839763
Q336	Hipoplasia e displasia do pulmão	4	4.44444	0.0741839763
Q878	Outras síndromes com malformações congênitas especificadas, não classificadas em outra parte	4	4.44444	0.0741839763
Q211	Comunicação interatrial	3	3.33333	0.0556379822
Q793	Gastrosquise	3	3.33333	0.0556379822
Q218	Outras malformações congênitas dos septos cardíacos	2	2.22222	0.0370919881
Q790	Hérnia diafragmática congênita	2	2.22222	0.0370919881
Q251	Coartação da aorta	2	2.22222	0.0370919881
Q870	Síndromes com malformações congênitas afetando predominantemente o aspecto da face	2	2.22222	0.0370919881
Q248	Outras malformações congênitas especificadas do coração	2	2.22222	0.0370919881
Q601	Agênese bilateral do rim	2	2.22222	0.0370919881
Q614	Displasia renal	1	1.11111	0.0185459941

Figura 30 – Interface dos resultados do submenu Frequência no banco SIM.

4.4.4. FERRAMENTAS DISTINTAS NO SINASC

O quarto módulo, Ferramenta SINASC, contém outros três submenus, sendo eles:

- Fluxo dos Nascidos Vivos:** permite visualizar o fluxo dos nascidos vivos com AC, em três cenários; (1) casos cujas mães e nascimentos ocorreram no mesmo local especificado; (2) casos cujas mães residem no local especificado, mas o nascimento foi realizado em outra cidade; (3) casos cujas mães não residem no local especificado, mas o nascimento foi realizado neste local. Para efetuar o fluxo, é preciso selecionar o município (local especificado), o período e o CID da AC (por Grupo, Categoria ou Subcategoria). Ver exemplo nas Figuras 31 e 32;
- Incidência:** permite visualizar o cálculo da incidência de notificação de AC tanto por local de residência da mãe como por local de ocorrência de nascimento. Pode-se escolher o estado e/ou município, o período e o CID da AC (por Grupo, Categoria ou Subcategoria). Ver exemplo nas Figuras 33 e 34;
- Inconsistência:** permite visualizar o cálculo da inconsistência das variáveis correlacionadas que notificam a AC. Os filtros permitem escolher o estado e/ou município e o ano do estudo. Ver exemplo nas Figuras 35 e 36.

Análise - Fluxo dos Nascidos Vivos com Anomalia Congênita

A ferramenta provê o Fluxo dos Nascidos Vivos com Anomalia Congênita. São apresentados ao usuário 3 cenários:

- (1) Apresenta a quantidade de casos cujas mães moram no lugar especificado e os nascimentos ocorreram no lugar especificado.
- (2) Apresenta a quantidade de casos cujas mães moram no lugar especificado mas os nascimentos ocorreram em outras cidades.
- (3) Apresenta a quantidade de casos cujas mães não moram no lugar especificado mas os nascimentos ocorreram no lugar especificado.

Assim, é permitido abaixo selecionar sua consulta:

- (a) em nível municipal;
- (b) em todo o período estudado ou em um ano específico e
- (c) envolvendo todas as anomalias congênicas ou uma anomalia específica.

Obs.: As especificações selecionadas serão consideradas no cálculo para que o resultado retrate a situação escolhida.

Localidade:

Estado:

Município:

Ano de Nascimento:

CID - Anomalia Congênita:

Grupo:

Categoria:

Subcategoria:

Figura 31 – Interface do submenu Fluxo dos Nascidos Vivos no banco SINASC.

Figura 32 – Interface dos resultados do submenu Fluxo dos Nascidos Vivos no banco SINASC.

Análise - Incidência de Anomalia Congênita

A ferramenta provê a Incidência de Anomalia Congênita levando em consideração a residência da mãe e a Incidência de Anomalia Congênita levando em consideração o local de ocorrência.

A Incidência de Anomalia Congênita por residência da mãe leva em conta o número de casos com anomalia congênita cujas mães vivem no local especificado e no ano selecionado dividido pelo total de nascidos vivos cujas mães vivem no local especificado e no ano selecionado. Essa medida nos permite levantar hipóteses sobre os fatores de risco envolvidos.

A Incidência de Anomalia Congênita por local de ocorrência leva em conta o número de casos com anomalia congênita que nasceram no local especificado e no ano selecionado dividido pelo total de nascidos vivos no local especificado e no ano selecionado. Essa medida nos permite visualizar possíveis municípios de referência nessa situação.

Assim, é permitido abaixo selecionar sua consulta:

- (a) em nível municipal;
- (b) em todo o período estudado ou em um ano específico e
- (c) envolvendo todas as anomalias congênitas ou uma anomalia específica.

Obs.: As especificações selecionadas serão consideradas no cálculo para que o indicador retrate a situação escolhida.

Localidade:

Estado:

Município:

Ano de Nascimento:

CID - Anomalia Congênita:

Grupo:

Categoria:

Subcategoria:

Figura 33 – Interface do submenu Incidência no banco SINASC.

Análise - Incidência de Anomalia Congênita

- PESQUISA REALIZADA -

Estado: SÃO PAULO
Município de Referência: RIBEIRÃO PRETO
Ano de Nascimento: 2010
CID: TODOS !

<p style="text-align: center;">INCIDÊNCIA POR LOCAL DE RESIDÊNCIA DA MÃE [?]</p> <p style="text-align: center;"><i>(Por 100 nascidos vivos)</i></p> <p>Incidência = Nascidos com AC/Total de Nascidos = 121/8141 = 1.49</p> <p style="text-align: center;"> Download dos casos relacionados</p>	<p style="text-align: center;">INCIDÊNCIA POR LOCAL DE OCORRÊNCIA [?]</p> <p style="text-align: center;"><i>(Por 100 nascidos vivos)</i></p> <p>Incidência = Nascidos com AC/Total de Nascidos = 205/11247 = 1.82</p> <p style="text-align: center;"> Download dos casos relacionados</p>
---	---

Observação: Os bancos de dados do SINASC são organizados por estado, portanto, ao calcularmos os resultados podem haver perdas de casos interestaduais (casos que migram entre as fronteiras dos estados).

Figura 34 – Interface dos resultados do submenu Incidência no banco SINASC.

Inconsistência nos campos com Anomalia Congênita - SINASC

A ferramenta provê uma verificação de inconsistência para o campo 34, que pode ser vista na fórmula a seguir, onde ID corresponde a variável que identifica a presença (1), ausência (2) ou ignorado (9) e COD representa o código da anomalia sendo preenchido (+) ou não preenchido (-):

(*) Inconsistência = $\{(ID(1)eCOD(-) + ID(2)eCOD(+) + ID(9)eCOD(+)) / (ID(1) + ID(2) + ID(9))\}$.

Assim, é permitido abaixo selecionar sua consulta:

(a) em nível estadual;
 (b) em nível municipal e
 (c) em todo o período estudado ou em um ano específico.

Localidade:
 Estado:
 Município:
 Ano de Nascimento:

Figura 35 – Interface do submenu Inconsistência no banco SINASC.

Análise - Inconsistência em variáveis correlacionadas de Anomalia Congênita

- PESQUISA REALIZADA -

Estado: SÃO PAULO
 Município de Referência: RIBEIRÃO PRETO
 Ano de Nascimento: 2001 a 2011

INCONSISTENCIA [?]
(Por 10.000 nascidos vivos)

Inconsistência = $\{(ID(1)eCOD(-) + ID(2)eCOD(+) + ID(9)eCOD(+)) / (ID(1) + ID(2) + ID(9))\} = 31/83352 = 3.7191669066129$

[Download dos casos relacionados](#)

Observação: Os bancos de dados do SINASC são organizados por estado, portanto, ao calcularmos os resultados podem haver perdas de casos interestaduais (casos que migram entre as fronteiras dos estados).

Figura 36 – Interface dos resultados do submenu Inconsistência no banco SINASC.

4.4.5. FERRAMENTAS DISTINTAS NO SIM

O quinto módulo, Ferramenta SIM, contém outros três submenus, sendo eles:

- a) **Fluxo de Mortalidade:** permite visualizar o fluxo de mortalidade por AC, em três cenários; (1) casos cujos indivíduos moraram no lugar especificado e o desfecho ocorreu no mesmo local; (2) casos cujos indivíduos moraram no lugar especificado, mas o desfecho ocorreu em outra cidade; (3) casos cujos indivíduos não moraram no lugar especificado, mas o desfecho ocorreu no lugar especificado. Para efetuar o fluxo, é preciso selecionar o município (local

especificado), o período e o CID da AC (por Grupo, Categoria ou Subcategoria). Ver exemplo nas Figuras 37 e 38;

- b) **Coefficiente de Mortalidade:** permite visualizar os cálculos da mortalidade infantil, perinatal e neonatal de pacientes que faleceram para cada 1.000 nascidos vivos. Pode-se escolher o estado e/ou município, sexo, o ano do estudo e qual coeficiente quer calcular e obter o resultado. Ver exemplo nas Figuras 39 e 40;
- c) **Mortalidade Segundo Causas:** permite visualizar o cálculo da mortalidade segundo causas notificadas pelo capítulo XVII da CID-10. Os filtros permitem escolher o estado e/ou município, sexo, o ano do estudo e o CID da AC (por Grupo, Categoria ou Subcategoria). Ver exemplo nas Figuras 41 e 42.

Análise - Fluxo de Mortalidade por Anomalia Congênita

A ferramenta provê o Fluxo de Mortalidade por Anomalia Congênita. Dos quais, são apresentados ao usuário 3 cenários:

- (1) Apresenta a quantidade de casos cujos indivíduos moraram no lugar especificado e o desfecho ocorreu no mesmo local.
- (2) Apresenta a quantidade de casos cujos indivíduos moraram no lugar especificado mas o desfecho ocorreu em outra cidade.
- (3) Apresenta a quantidade de casos cujos indivíduos não moraram no lugar especificado mas o desfecho ocorreu no lugar especificado.

Assim, é permitido abaixo selecionar sua consulta:

- (a) em nível municipal;
- (b) em todo o período estudado ou em um ano específico e
- (c) envolvendo todas as anomalias congênicas ou uma anomalia específica.

Obs.: As especificações selecionadas serão consideradas no cálculo para que o resultado retrate a situação escolhida.

Localidade:

Estado:

Município:

Ano de Nascimento:

CID - Anomalia Congênita:

Grupo:

Categoria:

Subcategoria:

Figura 37 – Interface do submenu Fluxo de Mortalidade no banco SIM.

Figura 38 – Interface dos resultados do submenu Fluxo de Mortalidade no banco SIM.

Coeficiente de Mortalidade - SIM

A ferramenta provê os dados dos coeficiente de mortalidade infantil, perinatal e neonatal de pacientes que vieram a óbito para cada 1.000 nascidos vivos numa determinada população e período. Assim, é permitido abaixo selecionar sua consulta:

- (a) em nível estadual;
- (b) em nível municipal;
- (c) por sexo e
- (d) por data específica.

Localidade:
 Estado:
 Município:

Sexo:
☐ Feminino ☐ Masculino ☒ Ambos

Mortalidade:
☐ Neonatal ☐ Perinatal ☒ Infantil

Data do Óbito:

Figura 39 – Interface do submenu Coeficiente de Mortalidade no banco SIM.

Resultado - Coeficiente de Mortalidade - SIM

- PESQUISA REALIZADA -
 Estado: SÃO PAULO
 Município de Referência: RIBEIRÃO PRETO
 Ano de Óbito: 2010

COEFICIENTE POR LOCAL DE RESIDÊNCIA [?] (Por 1000)	COEFICIENTE POR LOCAL DE OCORRÊNCIA [?] (Por 1000)
<p>CMI = Óbitos de crianças menores de um ano/Numero de Nascidos Vivos = (78/8141)*1000 = 9.58</p> <p style="text-align: center;"> Download dos casos relacionados</p>	<p>CMI = Óbitos de crianças menores de um ano/Numero de Nascidos Vivos = (185/11247)*1000 = 16.45</p> <p style="text-align: center;"> Download dos casos relacionados</p>

Observação: O banco de dados do SIM é organizado por estado, portanto, ao calcularmos os resultados podem haver perdas de casos interestaduais (casos que migram entre as fronteiras dos estados).

Figura 40 – Interface dos resultados do submenu Coeficiente de Mortalidade no banco SIM.

Mortalidade Segundo Causas - SIM

A ferramenta de mortalidade segundo causas mede taxa ou proporção de um determinado capítulo da CID - 10, por grupo de doença ou doença específica e que nesta ferramenta abordaremos crianças menores de um ano com notificação de Anomalia Congênita. Assim, é permitido abaixo selecionar sua consulta:

- (a) em nível estadual;
- (b) em nível municipal;
- (c) por sexo;
- (d) por ano específico e
- (e) envolvendo todas as anomalias congênicas por grupo, categoria ou uma anomalia específica.

Localidade:
 Estado:
 ▼
 Município:
 ▼

Sexo:
☐ Feminino ☐ Masculino ☒ Ambos

Ano do Óbito:
 ▼

CID - Anomalia Congênita:
 Grupo:
 ▼
 Categoria:
 ▼
 Subcategoria:
 ▼

Figura 41 – Interface do submenu Mortalidade Segundo Causas no banco SIM.

Resultado - Mortalidade Segundo Causas - SIM

- PESQUISA REALIZADA -
 Estado: SÃO PAULO
 Município de Referência: RIBEIRÃO PRETO
 Ano de Nascimento: 2010
 CID: TODOS !
 Sexo: AMBOS

TAXA POR LOCAL DE RESIDENCIA [?] (Por 1000)	TAXA POR LOCAL DE OCORRÊNCIA [?] (Por 1000)
<p>TMC = Taxa de óbito por AC em área e período escolhido/População desta área e período escolhido = $(18/8141) * 1000 = 2.21$</p> <p style="text-align: center;"> Download dos casos relacionados</p>	<p>TMC = Taxa de óbito por AC em área e período escolhido/População desta área e período escolhido = $(64/11247) * 1000 = 5.69$</p> <p style="text-align: center;"> Download dos casos relacionados</p>

Observação: O banco de dados do SIM é organizado por estado, portanto, ao calcularmos os resultados podem haver perdas de casos interestaduais (casos que migram entre as fronteiras dos estados).

Figura 42 – Interface dos resultados do submenu Mortalidade Segundo Causas no banco SIM.

4.4.6. MÓDULO LINKS

O sexto módulo, Links, permite acessar algumas referências citadas neste projeto, que podem ser vistas na Figura 43.

Figura 43 – Interface do menu Links.

4.4.7. MÓDULO QUEM SOMOS

O sétimo módulo, Quem Somos, permite visualizar a equipe, entrar em contato via e-mail e responder ao questionário eletrônico da pesquisa de satisfação, como visto nas Figuras 44, 45 e 46.

Figura 44 – Interface do submenu Equipe.

CONTATO

Nome:

E-mail:

Assunto:

Comentário:

Figura 45 – Interface do submenu Contato.

Questionário sobre o Sistema

Instruções para participar da pesquisa de satisfação.

1) Percorra os menus do site (Início, Histórico, Ferramentas SIHAC, SINASC, SIM, Links e Quem Somos) para conhecer a abrangência dos conteúdos abordados.

2) Nas Ferramentas explore as subtelas de cada menu, realize as pesquisas através dos filtros e clique sempre em Buscar.

3) Finalizado os itens anteriores, você estará pronto para responder às perguntas a seguir:

Dados pessoais:

Sexo:

☐ Feminino ☐ Masculino ☒ Outro

Faixa Etária:

Até 20 anos ▼

Nível Escolaridade:

Analfabeto ▼

Local de Residência(opcional):

Estado:

▼

Município:

-- Escolha um estado -- ▼

Dados sobre o Sistema PMAC:

0 - Você tem experiência com computadores?

Muito Baixo ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☒ 5 Muito Alto

1 - Como você achou que foi o tempo de resposta do sistema (feedback)?

Muito Baixo ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☒ 5 Muito Alto

2 - Você acha que o sistema PMAC possui similaridade com outras ferramentas ou sites que você já tenha usado?

Pouco similar ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☒ 5 Muito similar

3 - Você conseguiu sair das ferramentas, voltar à outra página ou refazer uma tarefa?

Dificuldade em navegar ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☒ 5 Facilidade em navegar

4 - Você achou as ferramentas padronizadas, ou seja, elas possuem filtros de busca similares umas com as outras?

Pouco padronizado ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☒ 5 Muito padronizado

5 - Você errou muito?

Poucas vezes ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☒ 5 Muitas vezes

6 - Você achou o sistema flexível e eficiente no tema das anomalias congênitas?

Pouco eficiente ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☒ 5 Muito eficiente

7 - O que você achou da estética do site?

Muito ruim ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☒ 5 Muito boa

8 - Você achou a linguagem do sistema simples e de fácil entendimento?

Muito fácil ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☒ 5 Muito difícil

9 - Você conseguiu encontrar a ajuda no sistema?

Muito pouco ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☒ 5 Muitas vezes

10 - Avalie no geral sua satisfação com o site:

Muito insatisfeito ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☒ 5 Muito satisfeito

Deixe neste espaço suas reclamações, críticas, sugestões ou elogios:(opcional)

Figura 46 – Interface do submenu Pesquisa de Satisfação.

4.4.8. TESTE DE AVALIAÇÃO

A fase de teste foi dividida em dois grupos. No primeiro foram selecionados quatro profissionais da informática biomédica, sendo que três integrantes não conheciam o sistema. Um material foi enviado via e-mail explicando o teste e a avaliação.

Lembrando que o teste caixa-preta foi empregado nos requisitos elencados para verificar a usabilidade das funções. Após o teste, a técnica adotada foi a avaliação heurística, que tem por finalidade apresentar aos avaliadores as dez estratégias predefinidas por Nielsen e consequentemente julgar se o sistema (site) está de acordo com os princípios de usabilidade. O intuito de cada avaliador foi medir até que ponto o site se aproxima dos requisitos iniciais. O resultado pode ser visto na Tabela 18.

Tabela 18 – Avaliação heurística dos especialistas.

Avaliação Heurística	Avaliador 1	Avaliador 2	Avaliador 3	Avaliador 4
Visibilidade do estado do sistema.	0	0	0	0
Correspondência entre o sistema e o mundo real.	1	1	2	2
Controle e liberdade do usuário.	0	0	0	0
Consistência e padrões.	0	1	0	1
Prevenção de erros.	1	1	1	0
Reconhecimento em vez de lembrança.	1	2	0	0
Flexibilidade e eficiência de uso.	1	1	0	0
Design estético e minimalista.	0	2	0	0
Ajudar usuários a reconhecer, diagnosticar e se recuperar dos erros.	1	1	0	0
Suporte e documentação.	1	2	0	0

Os itens com pontuações mais altas foram: correspondência entre o sistema e o mundo real, reconhecimento em vez de lembrança, design estético e minimalista, suporte e documentação. Todos estes tiveram a pontuação dois, que informa que há um problema menor com correção baixa de prioridade. O avaliador dois foi o que mais pontuou. Os itens que não apresentaram nenhum voto de usabilidade foram: visibilidade do estado do sistema e controle e liberdade do usuário.

O segundo grupo a passar pelo teste e avaliação foram os próprios usuários (navegadores) do site. No intuito de medir a usabilidade pelo usuário sem interferência de um conhecimento prévio na área de informática, foi desenvolvida a ferramenta de Pesquisa de Satisfação, cujas questões eletrônicas correspondem aos itens de Nielsen com cinco graus na escala de avaliação. Foram selecionados aleatoriamente cinco resultados que são apresentados na Tabela 19. Para o usuário, os itens dois e seis da avaliação estão na questão dois do site, e os itens nove e dez estão na questão nove do portal.

Tabela 19 – Avaliação heurística dos usuários.

Avaliação Heurística	Usuário 1	Usuário 2	Usuário 3	Usuário 4	Usuário 5
Visibilidade do estado do sistema.	0	2	1	0	0
Correspondência entre o sistema e o mundo real.	1	1	1	2	1
Controle e liberdade do usuário.	0	0	1	1	1
Consistência e padrões.	0	0	0	1	1
Prevenção de erros.	3	1	0	1	1
Reconhecimento em vez de lembrança.	1	1	1	2	1
Flexibilidade e eficiência de uso.	1	0	1	1	1
Design estético e minimalista.	1	0	1	1	1
Ajudar usuários a reconhecer, diagnosticar e se recuperar dos erros.	1	2	1	1	2
Suporte e documentação.	1	2	1	1	2

A avaliação realizada pelo grupo de usuários demonstra certa similaridade entre eles. Os itens com maiores pontuações de nota foram prevenção de erros, visibilidade do estado do sistema, correspondência entre o sistema e o mundo real, reconhecimento em vez de lembrança, ajudar usuários a reconhecer, diagnosticar e se recuperar dos erros, suporte e documentação. Não houve item com todos sem pontuar. O primeiro usuário pontuou valor três, que significa problema grave, correção de alta prioridade. No entanto, no campo de sugestão não havia críticas ou observação.

5. **DISCUSSÃO**

As análises da auditoria se iniciam com a apresentação da evolução do banco, sendo que a composição dos campos nas tabelas das bases apontou um crescimento de 156%, 61% e 55% no SINASC, SIM e SIH respectivamente.

Apesar de as bases terem crescido estruturalmente, foi visto através da evolução quantitativa que a inserção de registros dos estados ao longo do período de estudo se manteve constante, ou seja, para a maioria dos estados o acréscimo de campos a serem preenchidos não necessariamente influenciou na variação da quantidade de registros. Estados que tiveram variação podem estar relacionados com mais ou menos casos de nascimento, internação e óbito.

Os resultados trazidos pela incompletude nos indicam que, no geral, os três bancos de dados apresentaram dados excelente e bom, embora outras classificações tenham aparecido. Com estas informações pode-se estabelecer certa garantia da qualidade quanto à completude dos campos, uma vez que se adotou estratégia na escolha das variáveis utilizadas nas ferramentas, principalmente na técnica do relacionamento probabilístico, selecionando aquelas com melhores classificações.

A inconsistência obteve a classificação excelente ao longo do período, fortalecendo a segurança da análise da correlação entre as variáveis do campo que identifica a anomalia congênita no SINASC.

As análises dos coeficientes permitiram dimensionar os casos de morbidade e mortalidade no país. O coeficiente de incidência de notificação de AC mostrou crescimento ao longo do período de estudo e, em nível nacional, este crescimento foi de 54% desde 2001 até 2011. Em contrapartida, o coeficiente de mortalidade infantil e a taxa de mortalidade perinatal apresentaram um decréscimo de 31,9% e 25% respectivamente no período. Já as taxas de mortalidade neonatal tardia e mortalidade segundo causas de AC se mantiveram constantes (a primeira fica com dois e três casos para cada mil registros de nascimento, e a segunda com dois casos para cada mil óbitos).

A ideia da rastreabilidade surgiu com o intuito de auxiliar gestores e profissionais da área da saúde a identificar onde há notificações de AC, seja por nascimento, internação, seja óbito e assim auxiliar nas tomadas de decisões beneficiando a saúde coletiva. Ao estudar a técnica, foi preciso conhecer as bases. Uma vez identificada a não existência de campo

unívoco, foi necessário buscar um conjunto de campos comuns entre as bases que poderiam trazer resultados e contribuição ao estudo.

Inicialmente pensou-se em buscar subnotificação de AC nas bases, ou seja, relacionar dois bancos sem usar o campo com o CID do capítulo XVII com todos os registros disponíveis. No entanto, como o conjunto de variáveis comuns entre as bases foi pequeno e por não haver outras variáveis a acrescentar na tentativa de identificar um indivíduo, tomou-se a decisão de se ter apenas os dados com AC e, a partir deles, rastrear os notificados.

Definida a estratégia, levou-se em consideração que a notificação de AC rara seria mais bem identificada e acompanhada – para estes casos a técnica foi bem sucedida. Entretanto, para notificação de AC não rara e frequente, a identificação do registro como indivíduo foi mais difícil.

Durante o desenvolvimento do projeto, buscaram-se maneiras de melhorar a estratégia. Uma delas foi relacionar o SIH com o SIH na tentativa de ter um indivíduo a partir de um conjunto de internações que este tenha realizado, o que posteriormente declinou-se desta tentativa uma vez que o resultado não proporcionou acréscimos relevantes na notificação de AC não rara e tornaria o processo mais longo. Outra tentativa encontrada na literatura seria usar a variável número DN na base do SIM (esta variável foi introduzida nos registros públicos em 2006), no relacionamento SIM-SINASC. Contudo este campo é subnotificado e foi classificado como péssimo, o que inviabilizou o relacionamento.

Através dos resultados dos linkages vistos nas tabelas de rastreabilidade, percebe-se que os indivíduos que nasceram, foram internados e ou faleceram numa mesma cidade foram mais bem relacionados do que os que migraram.

O relacionamento probabilístico trouxe resultados significativos, proporcionando análises mais abrangentes ligando estas bases. Contudo, para garantir um melhor sucesso na técnica, o governo deveria criar campos de identificação dos registros entre as bases e as tornar um campo obrigatório usando a mesma ideia do número DN na base SIM, e que esta fosse acrescentada no SIH, bem como variáveis número DO do SIM e o número AIH do SIH estar nas outras bases.

O portal passou por três protótipos de design diferentes desde quando o projeto foi iniciado. Alguns colaboradores auxiliaram e opinaram na versão que hoje está no ar e que se encontra ilustrada na seção anterior.

Os testes realizados e apresentados pelos integrantes do grupo da área da informática não demonstraram grandes dificuldades, bem como os relatos trazidos por eles. Um destes foi: “As ferramentas são compostas por seleções de filtros e poucos botões. Esta estratégia

simplifica na execução e facilita o entendimento”. Uma sugestão foi: “Não é prioridade, no entanto mensagens de notificações poderiam aparecer com mais frequência”. O grupo composto por usuários apresentou valores mais homogêneos nos critérios de pontuação de usabilidade.

No geral, o feedback da maioria dos usuários com relação aos requisitos de usabilidade do site foi positivo, entretanto considerações pontuais foram levantadas. Uma vez que este projeto está em constante aperfeiçoamento, estas considerações já foram registradas e soluções traçadas para melhorar ainda mais o desempenho do site.

6. CONCLUSÃO

A partir deste estudo de desenvolvimento do sistema de vigilância de anomalias congênitas, utilizando como base bancos de dados públicos de todo o território nacional, no período de janeiro de 2001 a dezembro de 2011, com conteúdo relevante referente às anomalias, conclui-se que o portal PMAC mostrou-se útil ao agregar num mesmo site informações de nascimentos, internações e óbitos permitindo acesso centralizado aos dados públicos, além de deixar como legado ferramentas de gestão de informação, com módulos de auditoria, incidência, coeficientes, taxas e fluxos, apresentando dados em nível estadual e municipal.

Realizaram-se análises e estudos retrospectivos nas bases SINASC, SIH e SIM, além de aplicar a técnica do linkage e formar uma nova base, contribuindo com o aperfeiçoamento e expansão das informações notificadas por registros com AC nestas bases.

Foi um desafio usar a técnica de relacionamento probabilístico com três bases públicas num período de estudo de onze anos, cobrindo todos os estados e com um conjunto pequeno de variáveis comuns entre as bases, tornando-se um fator limitante para aprofundar nas análises de rastreabilidade. Entretanto os resultados dos indivíduos que nasceram, foram internados e ou vieram a óbito, numa mesma cidade, apresentaram bom linkage.

As etapas de teste e avaliação trouxeram enorme contribuição com notas, sugestões, críticas e elogios. Todo este processo fortaleceu e ampliou a observação e melhoria do portal, para que este seja utilizado por diferentes perfis de usuário que busquem informações de AC. Ressalta-se que o portal é dinâmico, e seu conteúdo seguirá sendo atualizado pelos envolvidos e futuros envolvidos, da mesma maneira que serão incluídas novas ferramentas e discussões a fim de acrescentar cada vez mais matéria-prima no auxílio à gestão das informações públicas e vigilância dos indivíduos com AC no Brasil.

REFERÊNCIAS

BRASIL. Ministério da Saúde. Fundação Nacional de Saúde (FUNASA). **Manual de procedimentos do sistema de informações sobre nascidos vivos**. Brasília, DF, 2001a.

BRASIL. Ministério da Saúde. Fundação Nacional de Saúde (FUNASA). **Manual de procedimentos do sistema de informações sobre mortalidade**. Brasília, DF, 2001b.

BRASIL. Ministério da Saúde. Fundação Nacional de Saúde (FUNASA). **Manual do sistema de informações hospitalares do SUS (SIH/SUS)**. Brasília, DF, 2004a.

BRASIL. Ministério da Saúde. **Auditoria no SUS**: noções básicas sobre sistemas de informação. Brasília, DF, 2004b. (Série A. Normas e manuais técnicos). Disponível em: <http://bvsmis.saude.gov.br/bvs/publicacoes/auditoria_sus.pdf>. Acesso em: 15 jan. 2017.

BRASIL. Ministério da Saúde. DATASUS. **Estrutura do SINASC para o CD-ROM**. 2005. Disponível em: <ftp://ftp.datasus.gov.br/dissemin/publicos/SINASC/NOV/Docs/Estrutura_SINASC_para_CD.pdf>. Acesso em: 15 jan. 2017.

BRASIL. Ministério da Saúde. DATASUS. **Estrutura do SIM para o CD-ROM**. 2006. Disponível em: <ftp://ftp.datasus.gov.br/dissemin/publicos/SIM/CID10/Docs/Estrutura_SIM_para_CD.pdf>. Acesso em: 15 jan. 2017.

BRASIL. Ministério da Saúde. **Auditoria do SUS orientações básicas**. 2011. Disponível em: <http://sna.saude.gov.br/download/LivroAuditoriaSUS_14x21cm.pdf>. Acesso em: 15 jan. 2017.

BRASIL. Ministério da Saúde. DATASUS. **Disseminação de informações do Sistema de Informações Hospitalares (SIH)**. 2015. Disponível em: <ftp://ftp.saude.sp.gov.br/Gis/BI/DicionarioBDSUS_SIH_RDSP.pdf>. Acesso em: 15 jan. 2017.

BRASIL. DATASUS. Ministério da Saúde. **CID – 10**. 2016. Disponível em: <<http://www.datasus.gov.br/cid10/V2008/cid10.htm>>. Acesso em: 26 dez. 2016.

BRASIL. DATASUS. Ministério da Saúde. - **Taxa de Mortalidade Neonatal Tardia**. 2017a. Disponível em: <<http://tabnet.datasus.gov.br/cgi/idb2000/fqc03.htm>>. Acesso em: 15 jan. 2017.

BRASIL. DATASUS. Ministério da Saúde. - **Taxa de Mortalidade Perinatal**. 2017b. Disponível em: <<http://tabnet.datasus.gov.br/cgi/idb2000/fqc05.htm>>. Acesso em: 15 jan. 2017.

BRASIL. Ministério da Saúde. **DATASUS**. 2017c. Disponível em: <<http://www2.datasus.gov.br/DATASUS/index.php>>. Acesso em: 15 jan. 2017.

BRITO, V. R. S. et al. Malformações congênitas e fatores de risco materno em Campina Grande - Paraíba. **Rev. Rene**, Fortaleza, v. 11, n. 2, p. 27-36, abr./jun. 2010.

CASTILHA, M. E.; ORIOLI, G. A representação social da mãe acerca da criança mal formada. **Rev. Esc. Enferm. USP**, v. 33, n. 2, p. 148-156, 2000.

CAMARGO, K.; COELI, C. M. **OpenRecLink**. 2016. Disponível em: <<http://reclink.sou.rceforge.net/>>. Acesso em: 15 jan. 2017.

COELI, C. M.; CAMARGO, K. R. Avaliação de diferentes estratégias de bloqueio no relacionamento probabilístico de registros. **Rev. Bras. Epidemiol.**, v. 5, n. 2, p. 185-196, 2002. Disponível em: <<http://www.scielo.org/pdf/rbepid/v5n2/06.pdf>>. Acesso em: 15 jan. 2017.

COELI, C. M.; CAMARGO, K. R. **Manual RecLink III**. Rio de Janeiro: 2007.

COELI, C. M.; CAMARGO, K. R. **OpenRecLink: guia do usuário**. Rio de Janeiro: 2012.

ESTUDO COLABORATIVO LATINO-AMERICANO DE MALFORMAÇÕES CONGÊNITAS (ECLAMC). **Bem vindos à web do Estudo Colaborativo Latino-Americano de Malformações Congênitas**. 2016. Disponível em: <<http://www.eclamc.org/port/index.php>>. Acesso em: 20 dez. 2016.

FRANCO, L. J.; PASSOS, A. D. C. (Orgs.). **Fundamentos de epidemiologia**. 2. ed. Barueri: Manole, 2011.

HOROVITZ, D. D. G.; LLERENA JUNIOR, J. C.; MATTOS, R. A. Atenção aos defeitos congênitos no Brasil: panorama atual. **Cad. Saúde Pública**, Rio de Janeiro, v. 21, n. 4, p. 1055-1064, jul./ago. 2005. Disponível em: <<http://dx.doi.org/10.1590/S0102-311X2005000400008>>. Acesso em: 15 jan. 2017.

MERCHÁN-HAMANN, E.; TAUIL, P. L.; COSTA, M. P. Terminologia das medidas e indicadores em epidemiologia: subsídios para uma possível padronização da nomenclatura. **Inf. Epidemiol. SUS**, Brasília, v. 9, n. 4, p. 276-284, dez. 2000. Disponível em: <<http://dx.doi.org/10.5123/S0104-16732000000400006>>. Acesso em: 15 jan. 2017.

NICOLA, P. D. R. et al. A utilização da Internet na notificação dos defeitos congênitos na Declaração de Nascido Vivo em quatro maternidades públicas do Município de São Paulo, Brasil. **Cadernos de Saúde Pública**, Rio de Janeiro, v. 26, n. 7, p. 1383-1390, jul. 2010. Disponível em: <<http://dx.doi.org/10.1590/S0102-311X2010000700017>>. Acesso em: 15 jan. 2017.

NIELSEN, J. **10 Usability heuristics for user interface design**. 1995. Disponível em: <<https://www.nngroup.com/articles/ten-usability-heuristics/>>. Acesso em: 20 dez. 2016.

OLIVEIRA, F. C. C. et al. Defeitos congênitos - tópicos relevantes. **Gazeta Médica da Bahia**, Salvador, v. 76, p. S32-S39, 2006. Suplemento 3. Disponível em: <<http://www.gmbahia.ufba.br/ojs/index.php/gmbahia/article/viewFile/281/272>>. Acesso em: 15 jan. 2017.

PEREIRA, P. K. et al. Malformação congênita do bebê e risco de transtornos mentais maternos durante o período gravídico-puerperal: uma revisão sistemática. **Cad. Saúde Colet.**, Rio de Janeiro, v. 19, n. 1, p. 2-10, 2011.

PRESSMAN, R. S. **Engenharia de Software**. Porto Alegre: McGraw Hill, 2006.

ROMERO, D. E.; CUNHA, C. B. Avaliação da qualidade das variáveis epidemiológicas e demográficas do Sistema de Informações sobre Nascidos Vivos, 2002. **Cad. Saúde Pública**, Rio de Janeiro, v. 23, n. 3, p. 701-714, mar. 2007. Disponível em: <<http://dx.doi.org/10.1590/S0102-311X2007000300028>>. Acesso em: 15 jan. 2017.

SANTOS, R. S.; DIAS, I. M. V. Refletindo sobre a malformação congênita. **Rev. Bras. Enferm.**, Brasília, v. 58, n. 5, p. 592-596, set./out. 2005. Disponível em: <<http://dx.doi.org/10.1590/S0034-71672005000500017>>. Acesso em: 15 jan. 2017.

SÃO PAULO. Prefeitura do Município. **Manual de aperfeiçoamento no diagnóstico de anomalias congênitas**. 2012. Disponível em: <http://www.prefeitura.sp.gov.br/cidade/secretarias/upload/saude/arquivos/sinasc/SINASC_ManualAnomaliasCongenitas_2012.pdf>. Acesso em: 20 dez. 2016.

SCHULLER-FACCINI, L. et al. Avaliação de teratógenos na população brasileira. **Ciênc. Saúde Coletiva**, v. 7, n. 1, p. 65-71, 2002. Disponível em: <<http://www.scielo.org/pdf/csc/v7n1/a06v07n1.pdf>>. Acesso em: 15 jan. 2017.

THE PHP GROUP. **Manual do PHP**. 2017. Disponível em: <http://php.net/manual/pt_BR/index.php>. Acesso em: 15 jan. 2017.

THE POSTGRESQL GLOBAL DEVELOPMENT GROUP. **PostgreSQL**. 2017. Disponível em: <www.postgresql.org/>. Acesso em: 15 jan. 2017.

THE WORLD WIDE WEB CONSORTIUM (W3C). **HTML**. 2017. Disponível em: <<http://www.w3.org/html/>>. Acesso em: 15 jan. 2017.

VICTORA, C. G.; BARROS, F. C. Infant mortality due to perinatal causes in Brazil: trends, regional patterns and possible interventions. **Sao Paulo Med. J.**, São Paulo, v. 119, n. 1, p. 33-42, jan. 2001. Disponível em: <<http://dx.doi.org/10.1590/S1516-31802001000100009>>. Acesso em: 15 jan. 2017.

W3SCHOOLS. **CSS tutorial**. 2017a. Disponível em: <<http://www.w3schools.com/css/>>. Acesso em: 15 jan. 2017.

W3SCHOOLS. **JavaScript**. 2017b. Disponível em: <<http://www.w3schools.com/js/>>. Acesso em: 15 jan. 2017.

ANEXOS

ANEXO A: Fluxo de Distribuição do Documento Padrão da DN.

Obs: O “-----” é para quando não existir o DIRES (Diretoria Regional de Saúde) ou equivalentes.

Fonte: BRASIL, 2001a - pág. 28.

ANEXO B: Fluxo da Informação – DNV Partos Hospitalares.

Fonte: BRASIL, 2001a - pág. 29.

ANEXO Ca: Fluxo da Informação – DNV Partos Domiciliares.

Fonte: BRASIL, 2001a - pág. 30.

ANEXO Cb: Fluxo da Informação – DNV Partos Domiciliares.

Fonte: BRASIL, 2001a - pág. 31.

ANEXO D: Fluxo do Sistema AIH.

ANEXO E: Fluxo de Distribuição do Documento Padrão da DO.

Obs: O “-----” é para quando não existir o DIRES; * ou equivalente; ** inclui Instituto Médico Legal (IML) ou Serviço de Verificação de Óbitos (SVO).

Fonte: BRASIL, 2001b - pág. 31.

ANEXO F: Fluxo da Informação Óbitos Hospitalares.

Fonte: BRASIL, 2001b - pág. 32.

ANEXO G: Óbitos por causas naturais em localidades sem médicos.

Fonte: BRASIL, 2001b - pág. 33.

ANEXO Ha: Parecer Comitê de Ética, p – 1. **PARECER CONSUBSTANCIADO DO CEP****DADOS DO PROJETO DE PESQUISA**

Título da Pesquisa: Desenvolvimento e Implantação de um Sistema para a Gestão da Informação e Vigilância de Anomalias Congênitas no Brasil

Pesquisador: Ananda Amorim Vieira

Área Temática:

Versão: 1

CAAE: 45519915.7.0000.5440

Instituição Proponente: Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 1.120.062

Data da Relatoria: 27/05/2015

Apresentação do Projeto:

Projeto intitulado "Desenvolvimento e Implantação de um Sistema para a Gestão da Informação e Vigilância de Anomalias Congênitas no Brasil".

Trata-se de um projeto a ser desenvolvido no Programa de Pós-graduação em Saúde na Comunidade, nível mestrado, sendo a mestranda com formação em Informática Biomédica. O orientador é docente do Departamento de Medicina Social.

Objetivo da Pesquisa:

Objetivo Primário:

Desenvolver sistema de vigilância de anomalias congênitas utilizando como base bancos de dados disponíveis através do DATASUS (SINASC, SIH e SIM).

Objetivo Secundário:

- Automatizar tanto estudos retrospectivos bem como contribuir para a formação de uma base de dados nacional de anomalias congênitas, com uma infraestrutura permanente para a base de dados relacionada através de linkage probabilístico.

- Implementação de um sistema (software) com módulos para a gestão da informação com um módulo de auditoria de dados e um módulo de análise da epidemiologia das anomalias congênitas,

Endereço: CAMPUS UNIVERSITÁRIO

Bairro: MONTE ALEGRE

CEP: 14.048-900

UF: SP

Município: RIBEIRÃO PRETO

Telefone: (16)3602-2228

Fax: (16)3633-1144

E-mail: cep@hcrp.usp.br

ANEXO Hb: Parecer Comitê de Ética, p – 2

Continuação do Parecer: 1.120.062

capaz de mostrar estatísticas regionalizadas tanto da incidência e prevalência, como da assistência a esse agravo.- Estudar e estabelecer dentro do sistema a capacidade de rastreabilidade de pacientes que ao nascerem com AC sofreram internações ao longo dos anos devido a um agravo específico e/ou vieram a óbito. Assim, através desse sistema é possível um melhor aproveitamento das informações disponíveis, mapeando os nascidos vivos nas internações pela busca de CIDs ligados as anomalias congênitas.

Avaliação dos Riscos e Benefícios:

Riscos: Não se aplica.

Benefícios: Acreditamos que este projeto seja estruturante da atenção à saúde em Genética Clínica, minimize desigualdades regionais e agregue informações epidemiológicas, na medida em que o produto principal seria um sistema que fornecesse informações sobre a notificação de casos de anomalias congênitas e que também permitisse uma avaliação do trânsito de um agravo específico dentro do sistema de saúde. Isto disponibilizado via web permitiria que diversos grupos pelo Brasil pudessem avaliar regionalmente e nacionalmente estas questões.

Comentários e Considerações sobre a Pesquisa:

Título adequado, projeto bem fundamentado, com bibliografia pertinente e atual. No tópico sobre critérios de inclusão e exclusão, não ficou clara a discriminação dos tipos de dados que serão considerados para o estudo (inclusão) e os que não serão (exclusão). Nestes tópicos, sugere-se que sejam redigidos de forma mais clara, possibilitando o estabelecimento do perfil de dados a serem utilizados no estudo e os não utilizados. Vale lembrar que critérios de exclusão, são aquelas condições que retiraria o dado da pesquisa se este preenchesse os critérios de inclusão e não uma negativa dos critérios de inclusão. Considerando-se que nenhum estudo é isento de risco, sugerimos que sejam elencados os riscos do estudo, como por exemplo, dificuldade de acesso ao DATASUS ou insucesso da pesquisa por alguma outra razão.

Considerações sobre os Termos de apresentação obrigatória:

A pesquisadora solicita a dispensa da apresentação do TCLE e justifica que a solicitação se baseia na metodologia do estudo, que utilizará dados secundários e públicos, disponibilizados pelo DATASUS.

Recomendações:

Não se aplica.

Conclusões ou Pendências e Lista de Inadequações:

Diante do exposto e à luz da Resolução CNS 466/2012, o projeto de pesquisa, assim como a

Endereço: CAMPUS UNIVERSITÁRIO
Bairro: MONTE ALEGRE **CEP:** 14.048-900
UF: SP **Município:** RIBEIRÃO PRETO
Telefone: (16)3602-2228 **Fax:** (16)3633-1144 **E-mail:** cep@hcrp.usp.br

ANEXO Hc: Parecer Comitê de Ética, p – 3

Continuação do Parecer: 1.120.062

dispensa do Termo de Consentimento Livre e Esclarecido, podem ser enquadrados na categoria APROVADO.

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

Considerações Finais a critério do CEP:

Projeto Aprovado: Tendo em vista a legislação vigente, devem ser encaminhados ao CEP, relatórios parciais anuais referentes ao andamento da pesquisa e relatório final ao término do trabalho. Qualquer modificação do projeto original deve ser apresentada a este CEP em nova versão, de forma objetiva e com justificativas, para nova apreciação.

RIBEIRÃO PRETO, 23 de Junho de 2015

Assinado por:
MARCIA GUIMARÃES VILLANOVA
(Coordenador)

Endereço: CAMPUS UNIVERSITÁRIO
Bairro: MONTE ALEGRE **CEP:** 14.048-900
UF: SP **Município:** RIBEIRÃO PRETO
Telefone: (16)3602-2228 **Fax:** (16)3633-1144 **E-mail:** cep@hcrp.usp.br