

ARTIGO ORIGINAL



Análise da qualidade dos dados de pré-natal de gestantes atendidas em Unidades Básicas de Saúde do município de São Paulo entre 2012 e 2020

Analysis of the quality of prenatal data of pregnant women attended at Healthcare Services in the city of São Paulo between 2012 and 2020

Fernanda Ferreira Corrêa^I , Thaís Rangel Bousquet Carrilho^{II} , Eliana de Aquino Bonilha^{III,IV} , Victor Nahuel Keller^{II} , Tarcisio Cantos de Melo^I , Gilberto Kac^{II} , Carmen Simone Grilo Diniz^I

^IUniversidade de São Paulo, Faculdade de Saúde Pública – São Paulo (SP), Brasil.

^{II}Universidade Federal do Rio de Janeiro, Instituto de Nutrição Josué de Castro, Observatório de Epidemiologia Nutricional – Rio de Janeiro (RJ), Brasil.

^{III}Centro Universitário São Camilo – São Paulo (SP), Brasil.

^{IV}Universidade de São Paulo, Faculdade de Saúde Pública, Grupo de Estudos Gênero, Evidências e Saúde – São Paulo (SP), Brasil.

RESUMO

Objetivo: Analisar a qualidade dos dados coletados no acompanhamento pré-natal registrados no Sistema Integrado de Gestão da Assistência à Saúde (SIGA) da Secretaria Municipal de Saúde de São Paulo de 2012 a 2020. **Métodos:** Estudo descritivo utilizando dados do SIGA e as variáveis: altura materna (cm), peso (kg) medido ao longo da gestação, idade gestacional na consulta pré-natal, pressão arterial (em mmHg) sistólica (PAS) e diastólica (PAD), e índice de massa corporal (IMC) no início da gestação (até 8 semanas). A análise da qualidade foi realizada por meio do cálculo dos indicadores: percentual de incompletude e valores zero de todas as variáveis estudadas, percentual de valores implausíveis de estatura, peso, IMC; preferência por dígito terminal do peso e estatura, e normalidade das distribuições. **Resultados:** Base de dados de gestantes disponibilizada para análise incluiu 8.046.608 registros e 1.174.115 mulheres. O percentual de incompletude e valores zeros foi baixo (<1%) em todas as variáveis originais do sistema. Existe maior número de registros ao final da gestação. Para as quatro variáveis de interesse originais do banco de dados (peso, altura, PAS, PAD), existe clara preferência por dígito terminal. As variáveis de interesse não apresentaram distribuição aproximadamente normal durante o período avaliado. **Conclusão:** A análise da qualidade mostrou necessidade de melhoria na padronização da coleta e do registro das informações, no arredondamento das medidas e na necessidade de incentivar as gestantes a iniciar o pré-natal o quanto antes; por isso, é importante investir na qualidade do dado, por meio de recursos educativos para profissionais que atuam na assistência. **Palavras-chave:** Atenção à saúde. Sistemas de dados. Gestantes. Indicadores de qualidade em assistência à saúde.

AUTORA CORRESPONDENTE: Fernanda Ferreira Corrêa. Rua Vital de Melo, 405, casa 113 –Vila Amato, CEP: 18087-632, Sorocaba (SP), Brasil. E-mail: fernandaferrera@correa@gmail.com

CONFLITO DE INTERESSE: nada a declarar.

COMO CITAR ESSE ARTIGO: Corrêa FF, Carrilho TRB, Bonilha EA, Keller VN, Melo TC, Kac G, et al. Análise da qualidade dos dados de pré-natal de gestantes atendidas em Unidades Básicas de Saúde do município de São Paulo entre 2012 e 2020. Rev Bras Epidemiol. 2023; 26: e230051. <https://doi.org/10.1590/1980-549720230051.2>

Esse é um artigo aberto distribuído sob licença CC-BY 4.0, que permite cópia e redistribuição do material em qualquer formato e para qualquer fim desde que mantidos os créditos de autoria e de publicação original.

Recebido em: 05/07/2023

Revisado em: 04/09/2023

Aceito em: 14/09/2023



INTRODUÇÃO

O registro de dados administrativos coletados na assistência pré-natal e do parto pode gerar grande volume de informações sobre os serviços prestados. Esses dados, quando coletados, registrados e tratados adequadamente, podem subsidiar gestores nas tomadas de decisão relacionadas ao planejamento da assistência e vigilância em saúde^{1,2}.

A exploração de bases de dados secundários, de uso administrativo, principalmente oriundas de sistemas de informação em saúde também pode contribuir de forma substancial para o conhecimento de necessidades locais. Entretanto, a utilização dessas bases deve ser precedida do entendimento dos seus limites, tais como a falta de padronização para a coleta e erros decorrentes da falta de consistência na entrada de dados no sistema. Além disso, é necessário conhecer a cobertura populacional, para delimitar a sua representatividade. Problemas existem tanto na qualidade quanto no tipo de informações obtidas; portanto, avaliar a completude e a qualidade desses dados é necessário³.

No município de São Paulo, um sistema para agendamentos de consultas, controle de fila de espera, controle de números de consultas por unidade durante o acompanhamento pré-natal foi criado em 2004. O Sistema Integrado de Gestão da Assistência à Saúde (SIGA) auxilia gestores na organização do sistema de referência e contrarreferência das unidades básicas de saúde e hospitais que realizarão os partos. Também possibilita caracterizar as gestantes atendidas no município; entretanto, a sua utilização tem tido caráter apenas administrativo, e houve pouca ou nenhuma utilização dos dados para um diagnóstico epidemiológico.

A análise estatística de conjuntos de dados muitas vezes se depara com desafios como a incompletude, valores zero e a implausibilidade. A incompletude refere-se à ausência de dados em certas observações, o que pode comprometer a integridade das análises. Valores zero podem indicar diferentes cenários, desde a ausência real de ocorrências até questões de relato. A implausibilidade aborda valores que estão fora do intervalo possível ou que não fazem sentido no contexto do estudo. Para avaliar a dissimilaridade entre conjuntos de dados, utiliza-se o índice de dissimilaridade, que quantifica as diferenças entre padrões de dados⁴. Estudos⁵⁻⁸ mostram que inconsistências encontradas em bases de dados podem prejudicar a qualidade da informação disponibilizada e a avaliação integral dos indivíduos.

O SIGA foi implantado em outros municípios do Brasil, porém, até o momento, não existem análises de qualidade dos dados gerados no município de São Paulo. Dessa forma, este estudo teve como objetivo analisar a qualidade dos dados coletados do acompanhamento pré-natal e registrados no SIGA da Secretaria Municipal de Saúde de São Paulo, entre 2012 e 2020.

MÉTODOS

Trata-se de estudo descritivo utilizando os dados de gestantes registradas no SIGA entre 2012 e 2020. Os dados são coletados na rotina do acompanhamento pré-natal e inseridos no sistema pelos profissionais das Unidades Básicas de Saúde (UBS). De um modo geral, os funcionários administrativos digitam e coletam informações sobre identificação do usuário e agendamento de consultas, e os técnicos (médicos, enfermeiras, nutricionistas) coletam e digitam as demais informações referentes ao atendimento, por exemplo, peso, altura, pressão arterial, entre outras.

A população residente do município de São Paulo em 2020 foi estimada pela Fundação SEADE em 11.869.660 habitantes⁹. No Sistema Único de Saúde (SUS) do município, as unidades responsáveis pela atenção básica à saúde totalizaram 469¹⁰ no ano de 2020, com 85% de UBS ou Unidades Básicas de Saúde da Família e as demais, Unidades Básicas de Saúde/Assistência Médica Ambulatorial (UBS/AMA). Todas essas unidades alimentam o SIGA com dados sobre os atendimentos na atenção básica e ambulatórios especializados. Neste trabalho foram analisados os dados de um dos módulos do sistema, o SIGA/Mãe paulistana, em que são inseridos dados das gestantes que realizam o pré-natal nas UBS.

Entre as variáveis do SIGA, foram consideradas para esta análise: altura materna (cm), peso (kg) medido ao longo da gestação, idade gestacional (em semanas) na consulta de pré-natal, pressão arterial sistólica (PAS) e pressão arterial diastólica (PAD), em mmHg. Duas variáveis derivadas foram criadas a partir dessas: o peso medido no início da gestação (até 8 semanas), pois se trata de variável que permite o cálculo do ganho de peso gestacional, e o índice de massa corpórea (IMC) no início da gestação (até 8 semanas), calculada a partir da divisão do peso medido até 8 semanas pela altura materna (em metros) ao quadrado. O IMC é utilizado para diagnóstico do estado nutricional materno no início da gestação¹¹.

A análise da qualidade da base foi realizada por meio do cálculo dos seguintes indicadores: percentual de incompletude e valores zero, percentual de valores implausíveis, preferência por dígito terminal e normalidade das distribuições. Esses indicadores foram adaptados a partir de indicadores utilizados para avaliação da qualidade de dados antropométricos de crianças¹².

A incompletude foi definida como o percentual de informações não preenchidas ou valores zerados (denominados valores zero)¹³. O grau de incompletude foi definido conforme os pontos de corte propostos por Romero e Cunha¹⁴: incompletude abaixo de 5% foi considerado excelente; de 5 a 9,9%, bom; de 10 a 19,9%, regular; de 20 a 49,9%, ruim; e 50% ou mais, muito ruim.

Valores implausíveis foram definidos de maneira distinta para cada variável de interesse, e, sempre que possível, referências externas foram utilizadas. Para idade materna,

valores inferiores a 10 e superiores a 55 anos foram considerados implausíveis; utilizamos esse critério considerando o perfil da idade da mãe dos nascidos vivos no Brasil¹⁵.

Para altura e IMC foram considerados implausíveis valores abaixo de -5 escore Z e acima de 5 escore Z, e para peso, valores abaixo de -6 escore Z e acima de 6 escore Z¹⁶. Utilizaram-se as curvas da Organização Mundial da Saúde (OMS)¹⁷ para definição dos pontos de corte.

Para efeito de ponto de corte de distribuição dos escores e maior similaridade entre os valores de adultas e adolescentes, foi padronizada a mesma curva, considerando a idade de 19 anos para todas as gestantes adultas, uma vez que não há uma curva de distribuição para valores implausíveis para maiores de 19 anos.

Para PAS e PAD, foram considerados valores implausíveis escores Z fora do intervalo de -6 a 6 da distribuição amostral para cada trimestre da gestação.

A preferência por dígitos terminais foi avaliada por meio de gráficos e pelo cálculo do índice de dissimilaridade, obtido por meio da seguinte fórmula (Equação 1):

$$\text{Índice de dissimilaridade} = \frac{\sum abs(\text{percentual observado} - \text{percentual esperado})}{2} \quad (1)$$

O índice de dissimilaridade pode ser interpretado como o percentual de valores que precisaria ser redistribuído para que a distribuição dos dígitos finais da variável de interesse seja uniforme (isto é, sem arredondamentos). Valores acima de 20% indicam preferência por dígito terminal e arredondamento para 0 ou 5¹².

O coeficiente de curtose indica a espessura das caudas da distribuição em comparação com a distribuição normal. Uma distribuição normal tem o coeficiente de curtose igual a zero. Valores positivos de curtose indicam que as caudas da distribuição são mais curtas que as da distribuição normal, e valores negativos indicam caudas mais longas¹⁸. Todas as análises foram realizadas no R versão 4.1.0¹⁹.

Este estudo é parte do projeto de pesquisa intitulado "Como tornar as intervenções no parto e seus desfechos mais visíveis aos sistemas de informação?" (projeto com financiamento já aprovado na Chamada Ciência de Dados para a Saúde Materno Infantil CNPq/Fundação Bill e Melinda Gates/2020/2022), aprovado pelo Comitê de Ética em Pesquisa da Secretaria Municipal da Saúde de São Paulo, sob o número 4.829.5.

RESULTADOS

A base de dados do SIGA de gestantes disponibilizada para análise incluía 8.046.608 registros e 1.174.115 mulheres no período de 2012 a 2020. O número de registros variou por ano de acompanhamento, com aumento no número de consultas registradas no sistema a partir de 2015 e pequena redução em 2018. Esse número se manteve estável entre os meses e anos de 2018 a 2020 (Figura Suplementar 1).

O percentual de incompletude e valores zeros foi baixo (<1%) em todas as variáveis originais do sistema, classificado como excelente¹⁴. Maiores valores de incompletude foram observados para as variáveis peso e IMC no início da gestação (75,69%) no ano de 2012, considerados muito ruim¹⁴. Não foram observadas variações nesses percentuais ao longo dos anos. Para as variáveis peso e IMC no início da gestação, o percentual de incompletude foi elevado em todo o período, superando 60% até 2017. Uma tendência de redução do percentual de incompletude nessas duas variáveis ao longo do tempo foi observada (para ambas as variáveis, 75,69% em 2012 *versus* 49,93% em 2020). Todas as mulheres registradas apresentaram pelo menos um registro de peso e um de altura. O percentual de escores Z implausíveis também foi baixo em todas as variáveis de interesse (<0,5%), com relativa estabilidade ao longo do tempo (Tabela 1 e Figura Suplementar 2).

Em relação ao número de registros (consultas) de acordo com a idade gestacional, é possível observar que o maior volume ocorre ao final da gestação, especialmente no último trimestre (a partir de 27 semanas) (Figura Suplementar 3). Esse padrão se repete para todos os anos avaliados, e 3.885.165 (48,28%) registros ocorreram entre 27 e 40 semanas de gestação, no período entre 2012 e 2020.

Para peso, altura, PAS e PAD, existe clara preferência por dígito terminal, isto é, os valores medidos são claramente arredondados para 0 ou 5 (Figura 1). Esse padrão de arredondamento é confirmado pelos valores observados para o índice de dissimilaridade, que é superior a 20% para peso, PAS e PAD, em todos os anos avaliados (Tabela 2). Para altura materna, apesar de a frequência de valores 0 e 5 ser superior à frequência observada para outros dígitos terminais (Figura 1), os valores para o índice de dissimilaridade permaneceram próximos a 10% no período (Tabela 2).

As variáveis de interesse não apresentaram distribuição aproximadamente normal no período avaliado. De maneira geral, as variáveis apresentaram valores de desvio-padrão próximos de 1, sendo o menor valor observado para PAD em 2014 (0,752) e o maior para IMC no início da gestação, em 2020 (1,223). PAS e PAD apresentaram assimetria à esquerda (negativa) em todo o período, exceto para PAS em 2012. As demais variáveis apresentaram coeficiente de assimetria próximo de zero. Todas as variáveis apresentaram coeficiente de curtose acima de 0; o menor valor foi 2,95 (IMC no início da gestação) e o maior foi 17,85 (PAD), o que indica distribuição leptocúrtica (Tabela 3).

DISCUSSÃO

Estudos sobre a qualidade de dados em países em desenvolvimento são limitados²⁰. Este estudo foi o pioneiro a qualificar os dados coletados no acompanhamento pré-natal e registrados no SIGA. Não existe consenso relativo aos critérios de avaliação da qualidade dos dados²¹. Nes-

Tabela 1. Frequência absoluta e relativa de valores de incompletude, valores zero e escores Z implausíveis nas variáveis de interesse do módulo de gestantes do Sistema Integrado de Gestão da Assistência à Saúde da Secretaria Municipal de Saúde de São Paulo, município de São Paulo, 2012 a 2020.

Variável*	Total	2012	2013	2014	2015	2016	2017	2018	2019	2020
Idade gestacional										
Total de registros	8.047.175	795.244	796.098	831.501	865.680	940.140	978.259	967.233	950.203	601.716
Incompletude	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Valores zero	567 (0,007)	27 (0,0034)	53 (0,0066)	58 (0,0072)	48 (0,0056)	66 (0,0072)	62 (0,0062)	100 (0,0104)	79 (0,0082)	74 (0,0077)
Altura materna										
Total de registros	8.047.175	795.244	796.098	831.501	865.680	940.140	978.259	967.233	950.203	601.716
Incompletude	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Valores zero	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Escores Z implausíveis	1.885 (0,023)	165 (0,021)	197 (0,025)	136 (0,017)	207 (0,024)	198 (0,022)	254 (0,026)	214 (0,022)	288 (0,03)	49 (0,031)
Peso no início da gestação (até 8 semanas)										
Total de registros	1.174.088	121.510	122.745	125.659	124.753	125.971	130.760	127.892	121.951	107.284
Incompletude	793.239 (67,56)	91.969 (75,69)	92.249 (75,15)	92.141 (73,33)	88.484 (70,93)	81.344 (64,57)	80.310 (61,42)	77.461 (60,57)	71.483 (58,62)	53.563 (49,93)
Valores zero	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Escores Z implausíveis	33 (0,003)	2 (0,007)	0 (0)	2 (0,006)	4 (0,011)	1 (0,002)	5 (0,01)	3 (0,006)	9 (0,018)	0 (0)
IMC no início da gestação										
Total de registros	1.174.088	121.510	122.745	125.659	124.753	125.971	130.760	127.892	121.951	107.284
Incompletude	793.247 (67,56)	91.970 (75,69)	92.249 (75,15)	92.141 (73,33)	88.485 (70,93)	81.345 (64,57)	80.312 (61,42)	77.464 (60,57)	71.483 (58,62)	53.563 (49,93)
Valores zero	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Escores Z implausíveis	776 (0,066)	36 (0,12)	35 (0,114)	48 (0,143)	57 (0,157)	73 (0,162)	111 (0,219)	101 (0,201)	135 (0,268)	26 (0,321)
Peso na gestação (distribuição por semana gestacional)										
Total de registros	8.047.175	795.244	796.098	831.501	865.680	940.140	978.259	967.233	950.203	601.716
Incompletude	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Valores zero	5 (0,0001)	0 (0)	1 (0,0001)	3 (0,0004)	0 (0)	0 (0)	0 (0)	0 (0)	1 (0,0001)	0 (0)
Escores Z implausíveis	1.531 (0,019)	87 (0,011)	110 (0,014)	102 (0,013)	112 (0,013)	183 (0,02)	228 (0,023)	213 (0,022)	239 (0,025)	41 (0,026)
Pressão arterial diastólica										
Total de registros	8.047.175	795.244	796.098	831.501	865.680	940.140	978.259	967.233	950.203	601.716
Incompletude	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Valores zero	1.524 (0,0189)	41 (0,0051)	171 (0,0213)	328 (0,0405)	183 (0,0215)	115 (0,0126)	140 (0,0141)	206 (0,0215)	176 (0,0184)	164 (0,0171)
Escores Z implausíveis	802 (0,01)	15 (0,002)	42 (0,005)	77 (0,01)	79 (0,009)	75 (0,008)	109 (0,011)	115 (0,012)	152 (0,016)	23 (0,015)
Pressão arterial sistólica										
Total de registros	8.047.175	795.244	796.098	831.501	865.680	940.140	978.259	967.233	950.203	601.716
Incompletude	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Valores zero	772 (0,0096)	13 (0,0016)	115 (0,0144)	270 (0,0334)	126 (0,0148)	42 (0,0046)	26 (0,0026)	85 (0,0089)	41 (0,0043)	54 (0,0056)
Escores Z implausíveis	3.148 (0,039)	58 (0,007)	153 (0,019)	312 (0,039)	263 (0,031)	212 (0,023)	405 (0,041)	474 (0,049)	604 (0,063)	96 (0,061)

IMC: índice de massa corpórea; * demais variáveis no banco de dados não apresentavam incompletude ou zeros.

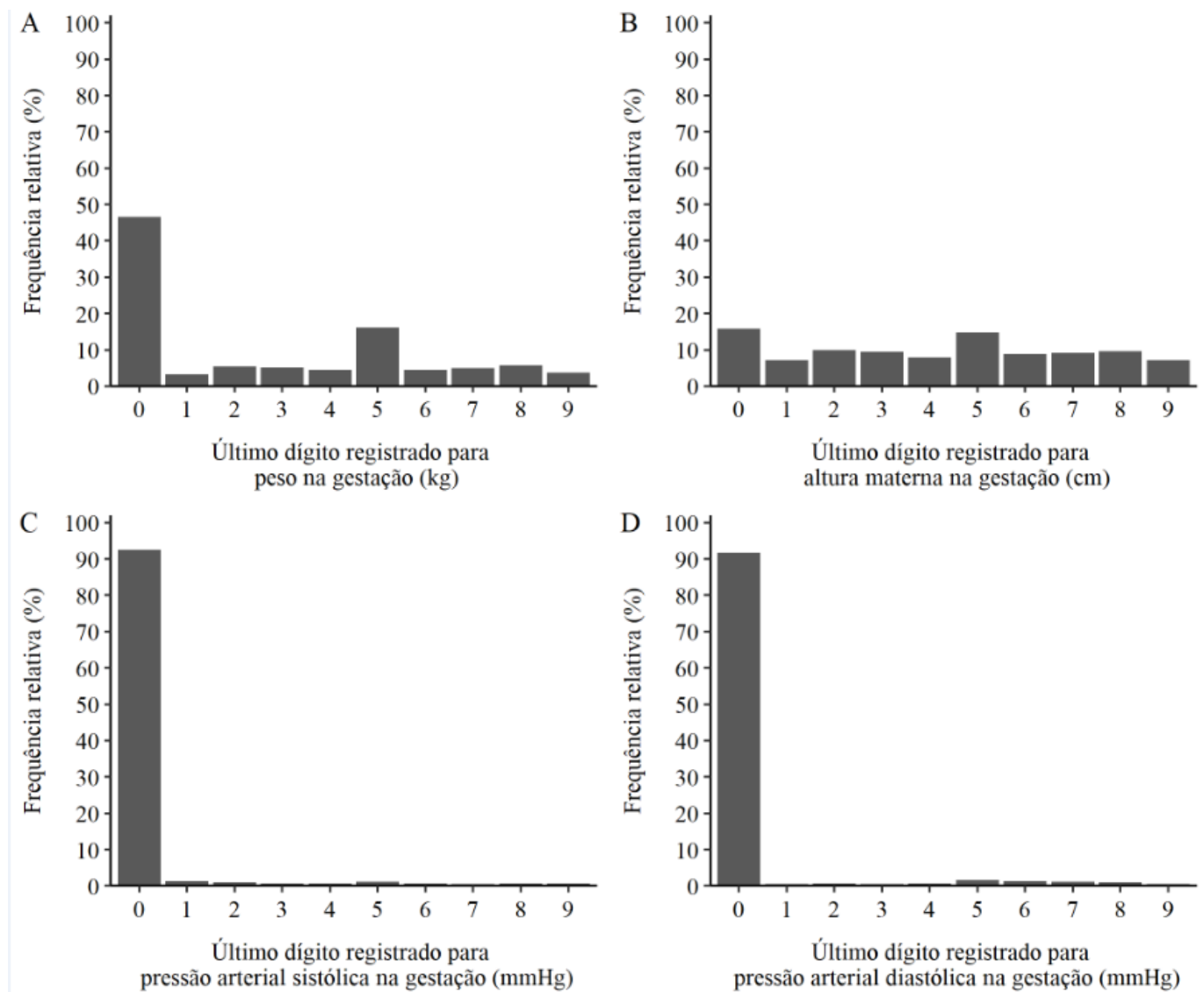


Figura 1. Identificação de preferência por dígito terminal nas variáveis antropométricas: (A) peso na gestação (kg); (B) altura materna (m); (C) pressão arterial sistólica (mmHg); e (D) pressão arterial diastólica (mmHg) do módulo de gestantes do Sistema Integrado de Gestão da Assistência à Saúde da Secretaria Municipal de Saúde de São Paulo, 2012 a 2020.

Tabela 2. Índice de dissimilaridade para as variáveis de interesse do módulo de gestantes do Sistema Integrado de Gestão da Assistência à Saúde, município de São Paulo, 2012 a 2020.

Variável de interesse	Total	2012	2013	2014	2015	2016	2017	2018	2019	2020
Peso na gestação	42,81	44,22	44,28	43,64	43,17	43,45	42,61	43,37	41,45	39,76
Altura materna	10,58	11,82	10,97	10,83	10,40	10,10	9,90	10,33	10,74	10,59
PAS	82,44	86,67	85,50	84,38	84,47	83,92	82,40	81,00	78,43	76,97
PAD	81,75	86,37	85,27	84,11	83,73	83,25	81,67	80,07	77,49	75,80

PAS: pressão arterial sistólica; PAD: pressão arterial diastólica.

te estudo, utilizamos percentual de incompletude e valores zero, percentual de valores implausíveis, preferência por dígito terminal e normalidade das distribuições.

Para o percentual de incompletude, valores classificados como muito ruim foram encontrados nas variáveis peso inicial e IMC inicial durante todo o período analisado. Valores de incompletude altos também foram encontra-

dos no estudo de Romero e Cunha¹⁴, que avaliou a qualidade da informação socioeconômica e demográfica, por Unidade Federada (UF) do Sistema de Informações sobre Mortalidade (SIM). Contudo, o percentual de incompletude e valores zeros foi baixo em todas as variáveis originais do sistema, considerando, nesse quesito, os dados de boa qualidade. Um banco de dados deve se mostrar completo

Tabela 3. Desvio-padrão, medidas de assimetria e curtose de escores Z das variáveis antropométricas de gestantes do município de São Paulo, 2012 a 2020.

Variável	Estatística	Total	2012	2013	2014	2015	2016	2017	2018	2019	2020
Altura materna	Desvio-padrão	0,996	0,986	0,983	0,989	0,990	0,990	0,998	1,001	1,006	1,002
	Assimetria	0,11	0,12	0,11	0,12	0,12	0,10	0,10	0,12	0,12	0,10
	Curtose	3,28	3,35	3,34	3,30	3,30	3,25	3,24	3,29	3,27	3,23
Peso no início da gestação (até 8 semanas)	Desvio-padrão	0,988	0,907	0,929	0,935	0,947	0,969	0,991	0,997	1,022	1,052
	Assimetria	0,98	1,09	1,05	1,05	1,01	0,96	0,97	0,93	0,93	0,89
	Curtose	4,60	4,98	4,85	4,89	4,77	4,58	4,61	4,39	4,40	4,37
IMC no início da gestação	Desvio-padrão	1,196	1,145	1,163	1,165	1,174	1,188	1,191	1,203	1,213	1,223
	Assimetria	0,03	0,14	0,11	0,09	0,06	0,02	0,04	-0,01	-0,03	-0,08
	Curtose	2,95	3,02	2,97	3,01	2,99	2,95	2,97	2,96	2,95	2,95
Peso na gestação (distribuição por trimestre)	Desvio-padrão	0,944	0,921	0,923	0,927	0,932	0,933	0,948	0,945	0,954	0,964
	Assimetria	0,45	0,45	0,44	0,44	0,45	0,43	0,44	0,44	0,49	0,45
	Curtose	4,37	4,19	4,20	4,34	4,48	4,36	4,37	4,38	4,45	4,35
Peso na gestação (distribuição por semana gestacional)	Desvio-padrão	0,950	0,930	0,930	0,934	0,939	0,939	0,953	0,950	0,959	0,968
	Assimetria	0,45	0,45	0,43	0,44	0,44	0,43	0,44	0,44	0,48	0,45
	Curtose	4,37	4,14	4,19	4,29	4,48	4,37	4,39	4,44	4,43	4,35
Pressão arterial sistólica	Desvio-padrão	0,941	0,778	0,849	0,929	0,912	0,890	0,922	0,963	0,989	1,038
	Assimetria	-1,92	-0,71	-1,33	-1,86	-1,80	-1,64	-1,83	-2,00	-2,08	-2,22
	Curtose	13,01	9,71	11,84	12,94	13,01	12,68	12,92	13,08	12,99	12,75
Pressão arterial diastólica	Desvio-padrão	0,967	0,752	0,859	1,011	0,928	0,878	0,948	0,977	1,051	1,060
	Assimetria	-2,65	-1,33	-2,20	-2,70	-2,52	-2,39	-2,62	-2,67	-2,76	-2,80
	Curtose	17,10	14,49	17,27	16,38	17,32	17,85	17,32	17,11	15,95	16,01

IMC: índice de massa corpórea.

e confiável quanto a seus registros²². Quando esses dados são inconsistentes, a confiabilidade das informações é prejudicada e falsos diagnósticos sobre a situação de saúde podem ser construídos²³.

Com relação à preferência por dígito terminal, neste estudo, existe uma preferência para os dígitos terminais 0 ou 5 nas variáveis peso, altura, PAS e PAD, semelhante ao Estudo Nacional de Alimentação e Nutrição Infantil (ENANI) de 2019¹⁶, que encontrou preferência por dígito terminal para o peso e a estatura.

Foram encontrados baixos percentuais de valores implausíveis de escore Z em todas as variáveis estudadas (<0,5%), como pode ser observado na Figura Suplementar 2, ficando dentro do recomendado internacionalmente (1%)¹².

Segundo o Ministério da Saúde, a gestante deve realizar, no mínimo, seis consultas de pré-natal, sendo, preferencialmente, uma no primeiro trimestre, duas no segundo trimestre e três no terceiro trimestre da gestação²⁴. O presente estudo mostra um aumento do número de registros de consultas de pré-natal no decorrer do período, porém não é possível saber se de fato o número de consultas aumentou ou se foram feitos mais registros no sistema. O baixo percentual de registros de consultas no início da gestação é similar aos achados de Domingues et al.²⁵ e Kac et al.¹¹, que evidenciam uma elevada proporção de gestantes brasileiras iniciando o acompanhamento do pré-natal após 12 semanas.

A aferição de peso deve ser realizada em todas as consultas de pré-natal, e a altura deve ser aferida pelo menos na primeira consulta de gestantes²⁶. Nos dados do SIGA, observa-se que todas as mulheres têm pelo menos um registro de peso e um de altura durante a gestação. Entretanto, quando se avalia a informação de peso coletada no início da gestação (até 8 semanas), o percentual de incompletude é alto em todo o período avaliado. A ausência de registro de peso no início da gestação é refletida no IMC no período, e representa uma importante limitação para caracterização do estado nutricional desses indivíduos e utilização de indicadores recomendados pelo Ministério da Saúde, tais como o ganho de peso na gestação. Nesse caso, poderia-se recomendar o uso do peso na ocasião do diagnóstico da gravidez, desde que as gestantes fossem orientadas a registrar o peso nessa ocasião.

O IMC pré-gestacional e o ganho de peso durante a gestação estão relacionados com os desenvolvimentos fetal e neonatal, assim como com desfechos obstétricos²⁷. Pautado nisso, destaca-se a importância de se coletar informações de qualidade durante o período pré-natal, contribuindo para um monitoramento mais eficaz e a obtenção de indicadores futuros mais precisos e valiosos, visando melhorias nas políticas públicas com orientação nutricional específica para esse grupo, melhorando o acesso das gestantes a alimentos menos processados, com menor teor de gorduras, sódio e açúcares, assim como com maior quantidade de vitaminas e minerais^{28,29}.

Foram encontradas algumas limitações do sistema (SIGA), como a idade gestacional estar em semanas e não em dias, o que reduz a precisão desse dado, também como resultado de arredondamentos não padronizados³⁰. Outra limitação do sistema foi a falta de padronização da coleta. No município de São Paulo, o processo de coleta de dados do pré-natal, assim como outros dados, ocorre em mais de 469¹⁰ pontos e são digitados por profissionais responsáveis em cada unidade. A formação em serviço desses profissionais nem sempre atende às suas necessidades, resultando, algumas vezes, em dados inconsistentes, não padronizados e incompletos³¹. No caso do peso e da estatura das gestantes, tanto as técnicas de mensuração quanto o registro podem estar incorretos, daí a relevância da instrumentalização dos profissionais envolvidos e da devolutiva dos dados para análise nas próprias unidades responsáveis pela coleta e pela digitação.

Este estudo é o primeiro a usar dados do SIGA, o que o torna inédito e de extrema importância para o conhecimento de gestores sobre dados das maternidades, UBS do seu território. Apesar desse importante ponto positivo, o presente estudo trouxe algumas limitações, como a utilização de indicadores de qualidade adaptados daqueles utilizados para crianças, por não existir indicadores específicos para gestantes.

As análises feitas neste estudo permitiram mostrar a necessidade de uma melhoria na padronização da coleta e do registro das informações, por exemplo, estipular regras claras no arredondamento das medidas para maior precisão. Outro ponto importante é a necessidade de incentivar as gestantes a iniciar o pré-natal o quanto antes, serem devidamente informadas e participarem ativamente do monitoramento de seus indicadores pessoal de saúde (como peso, glicemia e pressão arterial), assim como monitorar e melhorar a qualidade desses dados nos registros dos serviços.

O aprimoramento dos dados da base do SIGA resultaria em uma melhoria na qualidade dos dados, possibilitando ampliar o uso para cálculo de indicadores, para monitorar políticas públicas direcionadas para gestantes. Nesse sentido, é de suma importância investir na qualidade do dado, por meio de recursos educativos para os preenchedores³². A disseminação regular e o uso das informações sobre pré-natal do SIGA devem ser estimulados, pois são valiosos para análises epidemiológicas, podendo colaborar para a melhoria da sua qualidade¹. Ressalta-se a importância da avaliação da qualidade da base do SIGA, pois pode ser vinculada, por exemplo, com a base de nascidos vivos, o SINASC, para análise dos desfechos em relação ao bebê, à mãe e à condição de parto.

REFERÊNCIAS

1. Lima CRA, Schramm JMA, Coeli CM, Silva MEM. Review of data quality dimensions and applied methods in the evaluation of health information systems. *Cad Saude Publica* 2009; 25(10): 2095-109. <https://doi.org/10.1590/s0102-311x2009001000002>
2. Organización Panamericana de la Salud. Indicadores de salud: elementos básicos para el análisis de la situación de salud. *Boletín Epidemiológico*; 2001; 22(4)
3. Belo KO, Drumond Jr M. Funcionamento da atenção primária e acesso à atenção especializada. Dados secundários: processo de construção, análise e triangulação. In: Castro CP, Campos GWS, Fernandes JA, editors. *Atenção Primária e Atenção Especializada no SUS: análise das redes de cuidado em grandes cidades brasileiras*. São Paulo: Hucitec; 2021. p. 24-5.
4. Cunha G, Martins MR, Sousa RS, Oliveira FF. *Estatística aplicada às ciências e tecnologias da saúde*. Lisboa: Lidel; 2007.
5. Borges RC, Santos SLV, Lucena FN, Bachion MM. Segurança da informação: realidades na atenção primária em uma metrópole brasileira. *J Health Inform* 2023; 15(Especial). <https://doi.org/10.59681/2175-4411.v15.iEspecial.2023.1111>
6. Araujo YB, Rezende LCM, Queiroga MMD, Santos SR. Sistemas de Informação em Saúde: inconsistências de informações no contexto da Atenção Primária. *J Health Inform*. 2016; 8: 164-70
7. Digiacomio M, Davidson PM, Taylor KP, Smith JS, Dimer L, Ali M, et al. Health information system linkage and coordination are critical for increasing access to secondary prevention in Aboriginal health: a qualitative study. *Qual Prim Care*. 2010;18(1):17-26. PMID: 20359409
8. Silva MGP, Almeida GRP, Cabral DD, Paixão B, Morais IS, Carraro LZ, et al. 11º Congresso Brasileiro de Epidemiologia-Epidemiologia, Democracia e Saúde: Conhecimentos e Ações para Equidade. In: *Projeto Sisvan-Quali: Avaliação da Qualidade dos Dados do Estado do Rio de Janeiro*; 2021
9. Governo do Estado de São Paulo. Fundação Sistema Estadual de Análise de Dados. Portal de Estatísticas do Estado de São Paulo. Pesquisas em andamento [Internet]. [acessado em 28 set. 2023]. Disponível em: <http://produtos.seade.gov.br/produtos/projpop/>
10. Prefeitura da Cidade de São Paulo. Estabelecimentos e serviços de saúde [Internet]. [acessado em 28 set. 2023]. Disponível em: https://www.prefeitura.sp.gov.br/cidade/secretarias/saude/epidemiologia_e_informacao/informacoes_assistenciais/index.php?p=30566
11. Kac G, Carilho TRB, Rasmussen KM, Reichenheim ME, Farias DR, Hutcheon JA, et al. Gestational weight gain charts: results from the Brazilian Maternal and Child Nutrition Consortium. *Am J Clin Nutr* 2021; 113(5): 1351-60. <https://doi.org/10.1093/ajcn/nqaa402>
12. World Health Organization. United Nations Children's Fund. Recommendations for data collection, analysis and reporting on anthropometric indicators in children under 5 years old. Geneva: World Health Organization; 2019
13. Romero DE, Cunha CB. Evaluation of quality of epidemiological and demographic variables in the Live Births Information System, 2002. *Cad Saude Publica* 2007; 23(3): 701-14. <https://doi.org/10.1590/S0102-311x2007000300028>

14. Romero DE, Cunha CB. Avaliação da qualidade das variáveis sócio-econômicas e demográficas dos óbitos de crianças menores de um ano registrados no Sistema de Informações sobre Mortalidade do Brasil (1996/2001). *Cad Saude Publica* 2006; 22(3): 673-81. <https://doi.org/10.1590/S0102-311X2006000300022>
15. Brasil. Ministério da Saúde. Coordenação-Geral de Informações e Análises Epidemiológicas. Secretaria de Vigilância em Saúde. Nascidos vivos – Brasil [Internet]. 2023 [acessado em 29 jun. 2023]. Disponível em: <http://tabnet.datasus.gov.br/cgi/tabcgi.exe?sinasc/cnv/nvuf.def>
16. Universidade Federal do Rio de Janeiro. Aspectos metodológicos: descrição geral do estudo 1: ENANI 2019 [Internet]. 2021 [acessado em 30 jun. 2023]. Rio de Janeiro: UFRJ; 2021. Disponível em: https://enani.nutricao.ufrj.br/wp-content/uploads/2021/08/Relatorio1_ENANI-2019_Aspectos-Metodolo%CC%81gicos.pdf
17. World Health Organization. WHO child growth standards: length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: methods and development. Geneva: WHO; 2006
18. Pearson K. "Das Fehlergesetz und Seine Verallgemeinerungen Durch Fechner und Pearson." *A Rejoinder*. *Biometrika* 1905; 4(1/2): 169-212
19. R Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2021.
20. Lim YMF, Yusof M, Sivasampu S. Assessing primary care data quality. *Int J Health Care Qual Assur* 2018; 31(3):203-13. <https://doi.org/10.1108/IJHCQA-08-2016-0111>
21. Thiru K, Hassey A, Sullivan F. Systematic review of scope and quality of electronic patient record data in primary care. *BMJ* 2003; 326(7398): 1070. <https://doi.org/10.1136/bmj.326.7398.1070>
22. Macente LB, Zandonade E. Avaliação da completude do sistema de informação sobre mortalidade por suicídio na região Sudeste, Brasil, no período de 1996 a 2007. *J Bras Psiquiatr* 2010; 59(3): 173-81. <https://doi.org/10.1590/S0047-20852010000300002>
23. Costa JMBS, Frias PG. Avaliação da completude das variáveis da declaração de óbitos de menores de um ano residentes em Pernambuco, 1997-2005. *Ciênc Saúde Coletiva* 2011; 16(suppl 1): 1267-74. <https://doi.org/10.1590/S1413-81232011000700059>
24. Brasil. Ministério da Saúde. Secretaria de Atenção à Saúde. Departamento de Ações Programáticas Estratégicas. Área Técnica de Saúde da Mulher. Pré-natal e puerpério: atenção qualificada e humanizada. Brasília: Ministério da Saúde; 2005.
25. Domingues RMSM, Hartz ZMA, Dias MAB, Leal MC. Adequacy of prenatal care in the National Health System in the city of Rio de Janeiro, Brazil. *Cad Saude Publica* 2012; 28(3): 425-37. <https://doi.org/10.1590/s0102-311x2012000300003>
26. Brasil. Ministério da Saúde. Secretaria de Atenção à Saúde. Departamento de Atenção Básica. Atenção ao pré-natal de baixo risco. Brasília: Ministério da Saúde; 2012
27. Rovira MC, Baró AM, Guiu NB, Mesa LT, Heredia CL, Pérez SP, et al. The influence of obesity and diet quality on fetal growth and perinatal outcome. *Nutr Hosp* 2022; 39(6):1205-11. <https://doi.org/10.20960/nh.04076>
28. Malta MB, Gomes CB, Barros AJD, Baraldi LG, Takito MY, Benício MHDA, et al. Effectiveness of an intervention focusing on diet and walking during pregnancy in the primary health care service. *Cad Saude Publica* 2021; 37(5): e00010320. <https://doi.org/10.1590/0102-311x00010320>
29. Gomes CB, Vasconcelos LG, Cintra RMGC, Dias LCGD, Carvalhaes MABL. Hábitos alimentares das gestantes brasileiras: revisão integrativa da literatura. *Ciênc Saúde Colet* 2019; 24(6): 2293-306. <https://doi.org/10.1590/1413-81232018246.14702017>
30. Diniz CSG, Reis-Queiroz J, Kawai CA, Queiroz MR, Bonilha EA, NiyDY, et al. Dias potenciais de gravidez perdidos: uma medida inovadora da idade gestacional. *Rev Saude Publica* 2020; 54: 88. <https://doi.org/10.11606/s1518-8787.2020054002098>
31. Drumond EF, Machado CJ, Vasconcelos MR, França E. Utilização de dados secundários do SIM, Sinasc e SIH na produção científica brasileira de 1990 a 2006. *Rev Bras Estud Popul* 2009;26(1):7-19. <https://doi.org/10.1590/S0102-30982009000100002>
32. Chaves MMP, Miranda JL. Sistemas de Informação em Saúde: desafios encontrados durante a operacionalização e compartilhamento de dados. *Revista Eletrônica Acervo Saúde* 2023; 23(3): e11712. <https://doi.org/10.25248/reas.e11712.2023>

ABSTRACT

Abstract: Objective: To analyze the quality of data collected during prenatal care recorded in the Integrated Health Care Management System (SIGA) of the Municipal Department of Health of São Paulo from 2012 to 2020. **Methods:** Descriptive study using SIGA data and the variables: maternal height (cm), weight (kg) measured throughout pregnancy, gestational age at prenatal consultation, systolic (SBP) and diastolic (DBP) blood pressure (in mmHg), and body mass index (BMI) at the beginning of pregnancy (up to 8 weeks). Quality analysis was carried out by calculating the indicators: percentage of incompleteness and zero values of all variables studied, percentage of implausible values for height, weight, BMI; preference for terminal digit of weight and height, and normality of distributions. **Results:** The database of pregnant women made available for analysis included 8,046,608 records and 1,174,115 women. The percentage of incompleteness and zero values was low (<1%) in all original variables of the system. There are more records at the end of pregnancy. For the four original variables of interest in the database (weight, height, SBP, DBP), there is a clear preference for the terminal digit. The variables of interest did not present an approximately normal distribution during the evaluated period. **Conclusion:** The quality analysis showed the need for improving the standardization of information collection and recording, the rounding of measurements and the need for encouraging pregnant women to start prenatal care as soon as possible, in such a way that it is important to invest in data quality, through educational resources for professionals who work in health care.

Keywords: Health care. Data systems. Pregnant women. Quality indicators, health care.

AGRADECIMENTOS: Este trabalho foi apoiado, total ou parcialmente, pela Bill & Melinda Fundação Gates [ID INV-027961]. De acordo com as condições de subvenção da Fundação, uma Licença Genérica Creative Commons Attribution 4.0 já foi atribuída à versão do manuscrito aceito pelo autor que pode surgir desta submissão.

CONTRIBUIÇÕES DOS AUTORES: Corrêa, F. F.: Análise formal, Conceituação, Escrita – primeira redação, Escrita – revisão e edição. Carrilho, T. R. B.: Análise formal, Conceituação, Escrita – revisão e edição, Metodologia. Bonilha, E. A.: Conceituação, Curadoria de dados, Escrita – revisão e edição. Keller, V. N.: Análise formal, Metodologia. Melo, T. C.: Análise formal, Metodologia. Kac, G.: Escrita – revisão e edição. Diniz, C. S. G.: Administração do projeto, Escrita – revisão e edição, Obtenção de financiamento.

FONTE DE FINANCIAMENTO: Fundação Bill e Melinda Gates (OPP1201939) e Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (443775/2018 e 445710/2020-9).