

Curso Intermediário (apoiado por *software* R) da Análise da Situação de Saúde aplicado a Emergências Sanitárias, com foco na COVID-19

Aula 02- Sistemas de Informação em Saúde para Covid-19

Apresentação

Em um mundo em que a Transformação Digital avança cada vez mais, Sistemas de Informação tem se tornado cada vez mais importantes para a condução de políticas públicas. As decisões que são tomadas em âmbito público precisam ser informadas por dados seguros, precisos e de boa qualidade.

Um país continental como o Brasil produz diariamente quantidades elevadíssimas de dados, na área da saúde e fora dela. Para informar uma decisão de maneira adequada são necessárias análises dos dados disponíveis. Para isso, a qualidade dos dados assim como suas informações devem ser analisados.

Nessa aula vamos falar um pouco sobre os Sistemas de Informação em Saúde do Sistema Único de Saúde. Nossos sistemas são difundidos por todo o país, desde as unidades de saúde, passando por municípios e estados, até serem consolidados pelo Ministério da Saúde. Além disso, os dados são disponibilizados de forma pública.

1. Sistemas de Informação em Saúde (SIS)

Um Sistema de Informação em Saúde (SIS) são estruturas capazes de coletar, armazenar, codificar, processar e divulgar dados voltados para o acompanhamento da saúde de populações (1). Nesse sentido, os SIS são ferramentas fundamentais para a Vigilância Epidemiológica e promovem a máxima “informação para decisão e ação”.

As informações em saúde são fundamentais para a tomada de decisão em serviços de saúde. Nesse sentido, é importante criar um arcabouço de definições que serão úteis para compreender os SIS. Para efeitos didáticos desse curso, usaremos as seguintes definições:

- **Dado:** é definido com um fato objetivo. Um dado pode ser um número, um símbolo, uma figura, livres de qualquer contexto ou interpretação (2). Nesse sentido, é importante destacar que os dados são os registros brutos, livres de qualquer tratamento estatístico que são produzidos pelos SIS (1).
- **Informação:** pode ser definida como o conhecimento produzido a partir de dados. Ou seja, a informação é a transformação do dado para o conhecimento que agregará valor ao negócio e permitirá uma tomada de decisão que poderá ser mais adequada à situação conforme a capacidade das informações analisadas puderem ser adequadas à descrição de um determinado evento (1,3).

Em um contexto de SIS, o dado passará por algumas etapas, desde sua coleta até a disponibilização. Nesse sentido, o processo de coleta de dados é um dos mais importantes. Se houver erros na coleta de dados, a complexidade para sua resolução pode ser complicada posteriormente e pode significar a perda de uma informação.

Existem diversos instrumentos de coleta de dados para os SIS e nem todos eles são informatizados. Por exemplo, as fichas de notificação que alimentam o Sistema Nacional de Agravos de Notificação (SINAN), podem ser preenchidas manualmente e depois inseridas no sistema.

Além disso, há diversos outros instrumentos de coleta de dados para os sistemas de informação, como a Autorização de Internação Hospitalar (AIH) para o Sistema de Informações Hospitalares (SIH) e o Boletim de Produção Ambulatorial (BPA) para o Sistema de Informações Ambulatoriais (SIA).

Em nosso curso, trabalharemos dois sistemas de informação, que atualmente são utilizados para a notificação de Síndrome Respiratória por Covid-19, que é o e-SUS Notifica, e Síndrome Respiratória Aguda Grave por Covid-19, que é o Sistema de Vigilância Epidemiológica da Gripe (SIVEP-Gripe). Ambos os sistemas trabalham com notificações dos agravos analisados.

Após a coleta e inclusão de dados no sistema de informação os dados são armazenados em bancos de dados, sobre os quais são vamos nos aprofundar do ponto de vista computacional. Cada dado coletado é armazenado no banco de dados em variáveis. As variáveis são os atributos dos bancos de dados que permitem classificar um dado.

Nesse sentido, as variáveis podem ser classificadas em três principais formatos: os caracteres que permitem classificar os dados qualitativos nominais ou ordinais, incluindo nome, sexo, faixa etária e escolaridade, por exemplo. Os dados numéricos, que permitem classificar os dados quantitativos, discretos ou contínuos, por exemplo, idade, peso, altura e temperatura. Além desses, há os dados em formato de data, que classificam os períodos em dias, meses, anos e podem incluir horas, minutos e segundos.

Para cada banco de dados, há um dicionário de dados, que explicará não só o formato dos dados, mas também sua origem e seus valores. Os SIS que estão disponíveis para o SUS possuem diferentes bancos de dados e, logo, possuem diferentes dicionários de dados, que estão disponíveis e são importantes para subsidiar a análise dos dados em saúde.

Assim, o primeiro passo para a análise de um sistema de informação em saúde é compreender a origem dos dados, suas regras e quais são os dados disponíveis para a análise. Muitos dos sistemas de informação estão disponíveis de forma pública e é importante saber como obter e analisar esses dados.

2. Disponibilidade e *download* dos dados de SIS do SUS

Os dados de Sistemas de Informação em Saúde do SUS são armazenados a nível municipal pelas secretarias municipais de saúde, a nível estadual pelas secretarias estaduais de saúde e a nível nacional pelo Ministério da Saúde, por meio do Departamento de Informática do SUS (DATASUS/MS) (4).

As informações de saúde são consideradas informações sensíveis, por isso apenas poucos profissionais possuem acesso completo aos bancos de dados dos SIS, restringindo-se apenas aqueles que trabalham com essas informações. Com o advento da Lei Geral Proteção de Dados Pessoais (Lei nº 13.709, de 14 de agosto de 2018), a restrição de dados pessoais e dados sensíveis foi padronizada, sendo que a divulgação desses dados só pode ser realizada com a autorização das pessoas a cujo dados pertencem (5).

Apesar disso, o DATASUS disponibiliza diversos conjuntos de dados dos SIS de forma aberta e anonimizada, ou seja, não inclui informações que permitam identificar os indivíduos. Os dados públicos contêm diversas informações que podem ser usadas para análises ecológicas dos territórios em espaços de serviços de saúde e espaços acadêmicos.

Por meio da Transferência de Arquivos (6), o DATASUS disponibiliza os dados em formato “dbf” para a tabulação por meio do TabWin (Tabulador de dados para o Windows), uma ferramenta de tabulação de dados própria do DATASUS. Há também pacotes de análise desses dados disponíveis para o R.

São disponibilizados os dados de base populacional do Instituto Brasileiro de Geografia e Estatística (IBGE), a base territorial para construção de mapas, dados do Sistema de Comunicação de Informação Hospitalar (CIH), Comunicação de Informação Hospitalar e Ambulatorial (CIHA), Conjunto Mínimo de Dados (CMD), Cadastro Nacional de Estabelecimentos de Saúde (CNES), Notificações de casos suspeitos de SCZ (RESP), Sistema de Informações Ambulatoriais do SUS (SIASUS), Sistema de Informações Hospitalares do SUS (SIHSUS), Sistema de Informações de Mortalidade (SIM), Sistema de Agravos de Notificação Compulsória (SINAN), Sistema de Informação de Nascidos Vivos (SINASC) e Sistema de Monitoramento e Avaliação do Pré-Natal, Parto e Puerpério e Criança (SISPRENATAL)(6).

Uma alternativa de download de dados públicos é o openDataSUS, que fornece 18 conjuntos de dados dos SIS de forma aberta por meio de arquivos

em formato “csv” e outras que podem ser consultadas ou baixadas por meio de API (*Application Programming Interface*).

Atualmente, o openDataSUS possui uma diversidade de dados sobre a pandemia por Covid-19, que inclui dados de notificações, contratações e internações pela doença. Nesse curso, vamos nos ater aos dados de notificações por SR e SRAG do e-SUS Notifica e do SIVEP-Gripe, que estão disponíveis por meio do openDataSUS.

3. Qualidade dos dados de Sistemas de Informação para a Vigilância em Saúde

Uma parte importante da análise de dados de saúde é a avaliação da qualidade dos dados de seus sistemas de informação. Apesar disso, é reconhecido que os sistemas de informação são apenas uma parte daquilo que é necessário para compor um sistema de vigilância em saúde, que é mais abrangente e possui um processo de avaliação maior e mais complexo (7).

Nesse sentido, há diversas metodologias de avaliação para os Sistemas de Vigilância em Saúde que incluem a avaliação da qualidade dos dados que são produzidos em espaços de serviços de saúde. A avaliação perpassa os sistemas de informação, que serão abordados nessa aula como suporte para a análise de dados em saúde.

O Centro de Controle e Prevenção de Doenças dos Estados Unidos (CDC/USA), possuem um Guia para a Avaliação de Sistemas de Vigilância em Saúde. O guia cita os atributos que devem ser observados nos sistemas de vigilância e inclui: simplicidade, flexibilidade, qualidade dos dados, aceitabilidade, sensibilidade, valor preditivo positivo, representatividade, oportunidade e estabilidade. A avaliação desses atributos permite identificar a eficiência e recursos que podem ser usados (8).

- Sensibilidade: se refere à estrutura e facilidade de operação de um sistema de vigilância. Os sistemas devem ser o mais simples possíveis para a operação dos usuários permitindo que os objetivos do sistema sejam alcançados.
- Flexibilidade: um sistema flexível pode se adaptar às mudanças nas necessidades de informação ou nas condições operacionais em pouco tempo, com pouca necessidade de realocação de recursos humanos e estruturais adicionais. Isso pode incluir o registro de novos eventos de saúde, se necessário, e a integração a outros sistemas eletrônicos, por exemplo.
- Qualidade dos dados: a qualidade dos dados reflete a integridade e a validade dos dados. Isso significa que para a qualidade dos dados, deve-se observar a completude desses dados, ou seja, a quantidade de campos que foram preenchidas ou não. Além disso, analisar se os dados são verdadeiros pode exigir verificação manual, em alguns casos. É

possível realizar testes lógicos também para analisar se os campos são válidos com base em outros campos. Por exemplo, é possível encontrar possíveis erros em campos de idade ou sexo com base no campo que registra a gravidez em um sistema.

- Aceitabilidade: a aceitabilidade reflete a disposição de pessoas e organizações em participar dos sistemas de vigilância em saúde. Esse aspecto engloba mais do que apenas a participação no preenchimento de dados para sistemas de informação e envolve participação na vigilância ativa e passiva dos sistemas de vigilância.
- Sensibilidade: esse atributo ocorre em dois níveis, o primeiro se refere a proporção de casos de uma determinada doença ou agravo que são identificados pelo sistema. A segunda inclui a capacidade que o sistema fornece para identificar surtos e epidemias.
- Valor preditivo positivo (VPP): esse atributo se refere à proporção de casos notificados que realmente têm o evento relacionado à vigilância. Em conjunto com a sensibilidade, o VPP permite identificar o qual bem o sistema está operando.
- Representatividade: esse atributo está relacionado à precisão com a qual o sistema é capaz de descrever o evento que está sendo monitorado. A avaliação desse atributo pode se dar por meio da comparação dos eventos relatados com todos os eventos reais. Embora nem sempre seja possível conhecer todos os eventos reais, é possível realizar a comparação conhecendo a população afetada, o curso clínico da doença, as práticas médicas prevalentes e outras fontes de dados.
- Oportunidade: reflete a velocidade entre as etapas de registro, análise e tomada de decisão em um sistema. Ser oportuno significa que o sistema é eficiente no fornecimento de informações em um tempo que permita a prevenção de novos casos e o controle dos casos existentes.
- Estabilidade: refere-se à confiabilidade (ou seja, a capacidade de coletar, gerenciar e fornecer dados adequadamente sem falhas) e à disponibilidade (a capacidade de estar operacional quando necessário) do sistema de vigilância em saúde pública.

4. Sistemas de Informação em Saúde para Covid-19

A Covid-19 é uma doença causada pelo vírus SARS-CoV-2 que pode se manifestar por meio de uma Síndrome Gripal leve, com sintomas semelhantes a uma gripe, e a Síndrome Respiratória Aguda Grave, que pode acarretar na necessidade de uma internação e pode levar à morte (9).

Nesse sentido, há dois SIS no Brasil que realizam os registros de casos de Covid-19, sendo o e-SUS Notifica para os casos leves e o SIVEP-Gripe para os casos graves, internações e mortalidade. As internações também são registradas no Sistema de Informação Hospitalar, mas a finalidade desse sistema é promover o pagamento e registrar os gastos com internações, sendo menos utilizado na vigilância epidemiológica. Para a mortalidade, as mortes também são obrigatoriamente registradas no Sistema de Informação sobre Mortalidade (SIM).

a. Síndrome Respiratória (e-SUS Notifica)

Os casos leves de Síndrome Gripal (SG) suspeitos e confirmados por Covid-19 são registrados por meio do e-SUS Notifica, um módulo do e-SUS que é direcionado para a vigilância epidemiológica, incorporando os dados registrados na ficha de notificação de SG por Covid-19.

O sistema foi lançado em 27 de março de 2020 para receber dados dos casos suspeitos e confirmados de Covid-19 no Brasil. O sistema passa por várias modificações para acrescentar melhorias de usabilidade e dados ao sistema conforme as necessidades de vigilância se modificam no país (10).

O sistema é gerenciado pelo Ministério da Saúde, a nível nacional, e o módulo de notificações e monitoramento de contatos é de responsabilidade da Secretaria de Vigilância em Saúde (SVS). Os dados públicos desse SIS são disponibilizados online por meio do openDataSUS(11) e são atualizados simultaneamente por meio da API disponibilizada pelo MS no site, no endereço <<https://opendatasus.saude.gov.br/dataset/casos-nacionais>>.

b. Síndrome Respiratória Aguda Grave (SIVEP-Gripe)

Desde 2009, com a pandemia de H1N1, o SIVEP-Gripe desenvolve a vigilância da Síndrome Respiratória Aguda Grave (SRAG). Desde então, o sistema tem sido responsável pelos registros de casos de influenza e outros vírus respiratórios (12).

Em 2020, a vigilância de casos graves de Covid-19 foi incorporada ao sistema como reflexo da pandemia iniciada em 2019. Além do registro desses casos, o SIVEP também incorpora dados de internação e de óbitos, que podem ser analisados de maneira oportuna por seguirem um fluxo de informação mais rápido em relação ao SIM (12).

Assim como o e-SUS Notifica, o SIVEP-Gripe disponibiliza as informações de SRAG por Covid-19 no openDataSUS, incluindo as informações sobre notificação, casos confirmados, mortalidade e internação. Os dados públicos são anonimizados e não permitem a identificação dos indivíduos. Os dados podem ser acessados por meio do link: <<https://opendatasus.saude.gov.br/dataset/bd-srag-2021>>.

c. Mortalidade (SIM)

O Sistema de Informação de Mortalidade é responsável por registrar as mortes ocorridas no Brasil por qualquer causa, em qualquer local e qualquer momento. Os dados são advindos da Declaração de Óbitos e são fundamentais para compreender o perfil de mortalidade no Brasil (13).

Para análise dos casos de Covid-19, devem ser selecionados os óbitos com causa de óbito classificadas com a CID 10 usando os códigos B34.2 (infecção pelo coronavírus de localização não especificada) + o marcador U07.1 (Covid-19, vírus identificado) ou B34.2 (infecção pelo coronavírus de localização não especificada) + o marcador U07.2 (Covid-19, vírus não identificado) (14).

Nesse curso não abordaremos o SIM, mas é importante destacar que o sistema tem informações importantes sobre mortalidade de forma geral assim como no contexto da Pandemia de Covid-19.

5. Referências Bibliográficas

1. Brasil, Ministério da Saúde. Guia de Vigilância em Saúde [Internet]. 2º ed. Vol. Único. Brasília: Ministério da Saúde; 2017. Disponível em: <http://portalarquivos.saude.gov.br/images/pdf/2017/outubro/06/Volume-Unico-2017.pdf>
2. Bandeira J, Alcantara W, Sobrinho A, Ávila T, Bittencourt I, Isotani S. Dados Abertos Conectados. In 2014.
3. Simões AM. O processo de produção e distribuição de informação enquanto conhecimento: algumas reflexões. Perspect Em Ciênc Informação [Internet]. 1996 [citado 28 de agosto de 2021];1(1). Disponível em: <https://periodicos.ufmg.br/index.php/pci/article/view/22310>
4. DATASUS – Ministério da Saúde [Internet]. [citado 28 de agosto de 2021]. Disponível em: <https://datasus.saude.gov.br/>
5. L13709 [Internet]. [citado 28 de agosto de 2021]. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm
6. Transferência de Arquivos – DATASUS [Internet]. [citado 28 de agosto de 2021]. Disponível em: <https://datasus.saude.gov.br/transferencia-de-arquivos/>
7. Guidelines for Evaluating Surveillance Systems [Internet]. [citado 28 de agosto de 2021]. Disponível em: <https://www.cdc.gov/mmwr/preview/mmwrhtml/00001769.htm>
8. Updated Guidelines for Evaluating Public Health Surveillance Systems [Internet]. [citado 28 de agosto de 2021]. Disponível em: <https://www.cdc.gov/mmwr/preview/mmwrhtml/rr5013a1.htm>
9. Velavan TP, Meyer CG. The COVID-19 epidemic. Trop Med Int Health. março de 2020;25(3):278–80.
10. e-SUS Notifica [Internet]. Ministério da Saúde. [citado 28 de agosto de 2021]. Disponível em: <https://www.gov.br/saude/pt-br/composicao/vigilancia-em-saude-svs/sistemas-de-informacao/e-sus-notifica>
11. Notificações de Síndrome Gripal - Open Data [Internet]. [citado 28 de agosto de 2021]. Disponível em: <https://opendatasus.saude.gov.br/dataset/casos-nacionais>
12. SRAG 2021 - Banco de Dados de Síndrome Respiratória Aguda Grave - incluindo dados da COVID-19 - Open Data [Internet]. [citado 28 de agosto de 2021]. Disponível em: <https://opendatasus.saude.gov.br/dataset/bd-srag-2021>
13. Sistema de Informação sobre Mortalidade (SIM) - Dados Abertos - Centrais de Conteúdos - DASNT - SVS/MS [Internet]. [citado 28 de agosto de 2021]. Disponível em: <https://opendatasus.saude.gov.br/dataset/sim-dados-abertos>

de 2021]. Disponível em: <http://svs.aids.gov.br/dantps/centrais-de-conteudos/dados-abertos/sim/>

14. Brasil, Ministério da Saúde. Codificação das causas de morte com menção da síndrome inflamatória multissistêmica pediátrica no contexto da doença causada pelo Coronavírus Covid-19. 2020.