

UNIVERSIDADE ESTADUAL DE MARINGÁ
CENTRO DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA DE PRODUÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

AGEU DE ARAUJO MACHADO

Análise da Subnotificação de acidentes de trabalho fatais no Brasil

Maringá
2021

AGEU DE ARAUJO MACHADO

Análise da Subnotificação de acidentes de trabalho fatais no Brasil

Dissertação apresentada ao Programa de Pós-graduação em Engenharia de Produção do Departamento de Engenharia de Produção, Centro de Tecnologia da Universidade Estadual de Maringá, como requisito parcial para obtenção do título de Mestre em Engenharia de Produção.

Área de concentração: Engenharia de Produção

Orientador: Prof. Dr. Edwin Vladimir Cardoza Galdamez

Coorientador: Prof. Dr. Rodrigo Clemente Thom de Souza

Maringá
2021

Dados Internacionais de Catalogação-na-Publicação (CIP)
(Biblioteca Central - UEM, Maringá - PR, Brasil)

M149a	<p>Machado, Ageu de Araujo Análise da subnotificação de acidentes de trabalho fatais no Brasil / Ageu de Araujo Machado. -- Maringá, PR, 2021. 63 f.: il. color., figs., tabs.</p> <p>Orientador: Prof. Dr. Edwin Vladimir Cardoza Galdamez. Coorientador: Prof. Dr. Rodrigo Clemente Thom de Souza. Dissertação (Mestrado) - Universidade Estadual de Maringá, Centro de Tecnologia, Departamento de Engenharia de Produção, Programa de Pós-Graduação em Engenharia de Produção, 2021.</p> <p>1. Acidente de trabalho - Subnotificação - Brasil. 2. Segurança do trabalho. 3. Empresas - Atuação. 4. Captura e Recaptura - Técnicas. I. Galdamez, Edwin Vladimir Cardoza , orient. II. Souza, Rodrigo Clemente Thom de , coorient. III. Universidade Estadual de Maringá. Centro de Tecnologia. Departamento de Engenharia de Produção. Programa de Pós-Graduação em Engenharia de Produção. IV. Título.</p> <p>CDD 23.ed. 620.8</p>
-------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

FOLHA DE APROVAÇÃO

AGEU DE ARAUJO MACHADO

Análise da Subnotificação de acidentes de trabalho fatais no Brasil

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Produção do Departamento de Engenharia de Produção, Centro de Tecnologia da Universidade Estadual de Maringá, como requisito parcial para obtenção do título de Mestre em Engenharia de Produção pela Banca Examinadora composta pelos membros:

BANCA EXAMINADORA

Prof. Dr. Edwin Vladimir Cardoza Galdamez
Universidade Estadual de Maringá – DEP/UEM

Prof. Dr. Rodrigo Clemente Thom de Souza
Universidade Federal do Paraná - UFPR

Prof. Dra. Gislaine Camila Lapasini Leal
Universidade Estadual de Maringá – DEP/UEM

Prof. Dr. Otávio José de Oliveira
Universidade Estadual Paulista – DDP/FEG/UNESP

Aprovada em: 26 de agosto de 2021.

Local da defesa: meet.google.com/cgn-caio-rtm, conforme, PORTARIA CAPES Nº 36, DE 19 DE MARÇO DE 2020 e Ofício Circular nº 10/2020-DAV/CAPES

DEDICATÓRIA

A minha família por todo apoio.

AGRADECIMENTO

Agradeço primeiramente a Deus por ter me conduzido e direcionado até esse momento.

Aos meus pais e meus irmãos que sempre estiveram ao meu lado e deram todo apoio na busca por conhecimento e capacitação profissional, pessoal e educacional.

Ao meu orientador Dr. Prof. Edwin por partilhar comigo o seu conhecimento e orientações desde a graduação até o término do mestrado.

Ao coorientador Dr. Prof. Rodrigo, pelo suporte e orientações em momentos importantes da pesquisa.

Aos colegas de turma e professores do Programa de Pós-graduação em Engenharia de Produção que compartilharam esta etapa da minha formação.

A todas as pessoas que de alguma forma contribuiu com este momento.

A todos fica o meu muito obrigado!

EPÍGRAFE

Cada vez que você faz uma opção
está transformando sua essência
em alguma coisa um pouco
diferente do que era antes.

C. S. Lewis

Análise da Subnotificação de acidentes de trabalho fatais no Brasil

RESUMO

A evolução dos ambientes organizacionais nem sempre tem apresentado reflexo na saúde e segurança dos trabalhadores, tornando um fator de grande importância para desenvolvimento de um ambiente que garanta a integridade e saúde do trabalhador. Acidentes de trabalhos ocasionam prejuízos aos trabalhadores, às organizações e ao governo, principalmente quando o acidente ocorrido se caracteriza como grave e, nos piores cenários, conduz até mesmo ao óbito do trabalhador. O receio dos trabalhadores de sofrer alguma repreensão por parte da empresa ou a incúria das empresas em relatarem os acidentes que acontecem em suas dependências, podem levar a subnotificações de registros de acidentes de trabalhos nos bancos de dados. O objetivo do presente trabalho é utilizar a técnicas de Captura e Recaptura para determinar a subnotificação de acidentes de trabalhos fatais nos banco de dados relacionado a saúde segurança de trabalho, durante o ano de 2019. Fez-se uso do método de Captura-Recaptura por meio dos estimadores de Lincoln-Petersen e Chapman para o estudo de acidentes de trabalhos fatais no território brasileiro no ano de 2019, mediante relacionamento dos casos entre as fontes de dados da CAT, SIM e SINAN, sendo o primeiro administrado pela previdência social e os outros dois pelo Sistema Único de Saúde. Os dados foram previamente tratados para padronização dos registros e posteriormente utilizados dentro dos estimadores para determinação da subnotificação. Os resultados mostram que o banco de dados da CAT em comparação com o SINAN, apresenta uma taxa de subnotificação de 19% para casos que ocorrem a notificação da CAT, por sinalização do SINAN. Para acidentes fatais com relação empregado-empregador, a subnotificação atinge 41% entre CAT e SINAN, e 65% dos casos não ocorrem a abertura de CAT. A subnotificação entre o banco de dados da CAT e SIM, atinge uma taxa de 53%, e 83% dos acidentes não foram notificados a previdência social. Os resultados evidenciam a subnotificação de acidentes de trabalhos fatais no Brasil, identificando a necessidade de atuação nas empresas para redução na perda de vida nos ambientes de trabalho.

Palavras-chave: Subnotificação. Acidente de Trabalho. Captura e Recaptura.

Analysis of Underreporting of Fatal Labor Accidents in Brazil

ABSTRACT

The evolution of organizational environments has not always reflected on the workers' health and safety, making it a very important factor for the development of an environment that guarantees the worker's integrity and health. Labor accidents cause damage to workers, organizations, and the government, especially when the accident is characterized as serious and, in the worst scenarios, even leads to the worker's death. The fear of workers of suffering some reprimand by the company or the negligence of companies in reporting accidents that happen on their premises, can lead to underreporting of labor accident records in the databases. The objective of this work is to use the Capture-Recapture technique to determine the underreporting of occupational accidents sliced in the database related to health and safety at work, during the year 2019. The Capture-Recapture method was used by through the estimators of Lincoln-Petersen and Chapman for the study of fatal work accidents in the Brazilian territory in the year 2019, by linking the cases between the data sources of CAT, SIM and SINAN, the first being administered by social security and the two others by the Unified Health System. The data were previously processed to standardize the records and later used within the estimators to determine underreporting cases. The results show that the CAT database, compared to SINAN, presents an underreporting rate of 19% for cases where CAT notification occurs, by SINAN signaling. For fatal accidents with an employee-employer ratio, underreporting reaches 41% between CAT and SINAN, and 65% of cases do not open the CAT. Underreporting between the CAT and SIM databases reaches a rate of 53%, and 83% of accidents were not notified to social security. The results show the underreporting of work-related accidents in Brazil, identifying the need for action in companies to reduce loss of life in work environments.

Keywords: Underreporting. Labor accident. Capture and Recapture.

LISTA DE FIGURAS

FIGURA 1 – VISÃO GERAL DA PESQUISA.....	7
FIGURA 2 - FLUXOGRAMA DA GERAÇÃO DE DADOS RELACIONADOS A ACIDENTES DE TRABALHOS FATAIS.....	13
FIGURA 3 - CLASSIFICAÇÃO DOS MODELOS CAPTURA E RECAPTURA	17
FIGURA 4 - DIAGRAMA DE VENN E TABELA DE CONTINGÊNCIA DE UM PROCESSO DE CAPTURA- RECAPTURA ENVOLVENDO DUAS FONTES.....	19
FIGURA 5 - VISÃO GERAL DO PROCESSO DE DESCOBRIMENTO DO CONHECIMENTO.....	21
FIGURA 6 - METODOLOGIA DE PESQUISA.....	26
FIGURA 7 – QUANTIDADE DE REGISTROS POR BASE DE DADOS	33
FIGURA 8 - DISTRIBUIÇÃO DOS REGISTROS POR FAIXA ETÁRIA, SEXO E BASE DE DADOS	35
FIGURA 9 - QUANTIDADE DE REGISTRO POR FAIXA ETÁRIA BASE DE DADOS	36
FIGURA 10 - DISTRIBUIÇÃO DOS REGISTROS POR UNIDADE FEDERADA	37
FIGURA 11 - DIAGRAMA DE VENN ENTRE CAT E SINAN	44
FIGURA 12 - DIAGRAMA DE VENN ENTRE CAT E SIM.....	46

LISTA DE TABELAS

TABELA 1 - CARACTERIZAÇÃO DOS BANCOS DE DADOS DE SSO NO BRASIL	31
TABELA 2 - PADRONIZAÇÃO DAS VARIÁVEIS DAS FONTES DE DADOS	32
TABELA 3. QUANTIDADE DE ACIDENTE POR SEXO E FONTE DE DADOS	34
TABELA 4 - QUANTIDADE DE ACIDENTES FATAIS NA CAT POR CLASSIFICAÇÃO DO CNAE.....	36
TABELA 5 - QUANTIDADE DE ACIDENTE FATAL ABSOLUTO E RELATIVO DA CAT	38
TABELA 6 - CLASSIFICAÇÃO DAS VARIÁVEIS DOS <i>DATASETS</i>	41
TABELA 7 - SUBNOTIFICAÇÃO DE ACIDENTES FATAIS CAT E SINAN.....	45
TABELA 8 - SUBNOTIFICAÇÃO DE ACIDENTES FATAIS CAT E SIM.....	46

LISTA DE ABREVIATURAS E SIGLAS

AT	<i>Acidente de Trabalho</i>
ATF	<i>Acidente de Trabalho Fatal</i>
BO	<i>Boletim de Ocorrência</i>
CAGED	<i>Cadastro Geral de Empregados e Desempregados</i>
CAT	<i>Comunicação de Acidente de Trabalho</i>
DATAPREV	<i>Empresa de Tecnologia e Informações da Previdência</i>
DATASUS	<i>Departamento de Informática do SUS</i>
DO	<i>Declaração de Óbito</i>
ILO	<i>International Labour Organization</i>
IML	<i>Instituto Médico Legal</i>
INSS	<i>Instituto Nacional do Seguro Social</i>
ISO	<i>International Organization for Standardization</i>
LAI	<i>Lei de Acesso à Informação</i>
LC	<i>laudo cadavérico</i>
OIT	<i>Organização Internacional do Trabalho</i>
RAIS	<i>Relação Anual de Informações Sociais</i>
SIH-SUS	<i>Sistema de Informações Hospitalares</i>
SIM	<i>Sistema de Informações sobre Mortalidade</i>
SINAN	<i>Sistema de Informação de Agravos de Notificação</i>
SISCAT	<i>Sistema de Catalogação Brasileiro</i>
SST	<i>Saúde e Segurança do Trabalho</i>
SUIBE	<i>Sistema Único de Informações de Benefícios</i>
SUS	<i>Sistema Único de Saúde</i>

SUMARIO

1	INTRODUÇÃO.....	1
1.1	OBJETIVOS DO ESTUDO	4
1.1.1	<i>Objetivos específicos</i>	5
1.2	JUSTIFICATIVA	5
1.3	ESTRUTURA DO TRABALHO.....	6
2	REVISÃO DE LITERATURA	9
2.1	SAÚDE E SEGURANÇA DO TRABALHO.....	9
2.1.1	<i>Notificação de acidente de trabalho</i>	12
2.1.2	<i>Registros de acidentes de trabalho fatais</i>	13
2.2	SUBNOTIFICAÇÃO DE ACIDENTES DE TRABALHO	15
2.2.1	<i>Captura e Recaptura</i>	17
2.3	SELEÇÃO E PRÉ-PROCESSAMENTO DE DADOS	21
2.4	TRABALHOS CORRELATOS.....	22
3	MÉTODO DE PESQUISA	25
3.1	CARACTERIZAÇÃO DA PESQUISA	25
3.2	MÉTODO DE ESTUDO.....	26
4	ANÁLISE EXPLORATÓRIA	29
4.1	DADOS DA SAÚDE E SEGURANÇA DO TRABALHO	29
4.2	PREPARAÇÃO DOS DADOS.....	32
4.3	ANÁLISE EXPLORATÓRIA DOS DADOS	33
5	ANÁLISE DE RESULTADOS	40
5.1	DETERMINAÇÃO DAS VARIÁVEIS.....	41
5.2	TRATAMENTO DOS DADOS.....	42
5.3	PREMISSAS.....	42
5.4	REQUISITOS CAPTURA E RECAPTURA.....	43
5.5	CAT E SINAN	44
5.6	CAT E SIM	45
5.7	DISCUSSÃO DOS RESULTADOS	47
5.7.1	<i>Subnotificação CAT x SINAN</i>	49
5.7.2	<i>Subnotificação CAT x SIM</i>	50

6	CONCLUSÃO	52
	6.1 CONSIDERAÇÕES FINAIS.....	52
	6.2 LIMITAÇÕES DO ESTUDO	54
	6.3 TRABALHOS FUTUROS.....	55
7	REFERÊNCIAS	56

INTRODUÇÃO

De acordo com a *International Labour Organization* - ILO (2018) acontece mais de 2,78 milhões de mortes por ano, como resultado de acidentes ou doenças relacionadas ao trabalho. Este número é decorrente dos aproximados 374 milhões de lesões não fatais relacionadas ao trabalho a cada ano, resultando em mais de 4 dias de ausência no trabalho. O custo humano dessa adversidade diária é vasto e o ônus econômico de práticas inadequadas de segurança e saúde ocupacional é estimado em 3,94% do Produto Interno Bruto global a cada ano (ILO, 2020). É fundamental portanto que as organizações considerem na rotina do gerenciamento das atividades de prevenção a proteção da saúde e segurança de seus trabalhadores durante as atividades (CIARAPICA; GIACCHETTA; 2009, KHOSROWABADI; GHOSI; MAKUI, 2019, SINGH; MISRA, 2020).

As constantes mudanças nos ambientes internos e externos das organizações as levam a busca pela otimização da produtividade do trabalhador, e juntamente com esta preocupação, a Saúde e Segurança do Trabalho também é de extrema importância do ponto de vista moral, jurídico e econômico (BEVILACQUA; CIARAPICA; GIACCHETTA, 2008, DA SILVA; AMARAL, 2019, SINGH; MISRA, 2020).

Embora a saúde e segurança do trabalho seja regulamentada por vários aspectos legislativos nacionais, a análise de acidentes de trabalho nem sempre é uma prioridade (SÁNCHEZ, 2011). A segurança do ambiente de trabalho pode ser melhorada a partir do desenvolvimento de diretrizes

que evitam a redução da probabilidade ou gravidade dos eventos inesperados de segurança (KHOSROWABADI; GHOUZI; MAKUI, 2019).

Os acidentes e as doenças de trabalho são um enorme problema de saúde pública (TEUFER *et al.*, 2019) que precisa de políticas públicas e intervenções para melhorar a Saúde e Segurança do Trabalho (RAMOS; AFONSO; RODRIGUES, 2020). Um fator que contribui para o problema é a presença de uma crença fatalista sobre a inevitabilidade de acidentes, ocasionando uma reação indesejada a esse pensamento no relaxamento dos esforços para reduzir a frequência de acidentes (SÁNCHEZ, 2011). Em geral, os acidentes de trabalho ocorrem devido à falta de conhecimento, treinamento, gerenciamento, meios para executar a tarefa com segurança ou, alternativamente, devido a um erro de julgamento, descuido, apatia ou imprudência total (LIAO; PERNG, 2008, SARKAR *et al.*, 2019). Acidentes de trabalho e lesões ocupacionais podem afetar a vida dos trabalhadores, dentro e fora do ambiente de trabalho, e impõem uma carga econômica considerável aos empregadores, trabalhadores, companhias de seguros, sistemas de assistência médica e sociedade (SÁNCHEZ, 2011, KAKHKI; FREEMAN; MOSHER, 2019).

Quando um acidente ocorre, o mesmo deve ser reportado aos órgãos responsáveis, podendo ser tanto pelo trabalhador, vítima do acidente, quanto pela empresa na qual está vinculado. O conjunto destes registros é conhecido como banco de dados de acidentes de trabalho (AHMED; SADULLAH; YAHYA, 2019).

No Brasil a comunicação do acidente ou aquisição de doença de trabalho por parte do trabalhador na Previdência Social ocorre por meio da Comunicação de Acidente de Trabalho (CAT). A CAT é a notificação que o Instituto Nacional de Seguro Social (INSS) realiza todas as garantias de seguro prevista em lei ao trabalhador. O Artigo 169 da Consolidação das Leis do Trabalho (CLT) impõe a obrigatoriedade da notificação das doenças profissionais e das produzidas por condições especiais de trabalho, comprovadas ou suspeitas, tornando crime pelo Artigo 269 do Código Penal caso não ocorra a denúncia à autoridade pública. A legislação não impõe prazo para o trabalhador fazer a notificação, contudo a empresa tem o prazo de um dia útil seguinte da ocorrência para comunicar o acidente de trabalho a Previdência Social e em caso de morte, de forma imediata, sob pena de multa conforme o Artigo 22 da Lei 8213/1991. Conforme a situação que o acidente acontece, o registro pode ser realizado em outras plataformas que indica a classificação com um acidente de trabalho, ou em alguns casos a apenas a notificação da morte relacionado a um acidente que ocorreu dentro do ambiente de trabalho. Dentre estas bases, destaca-

se o Sistema de Informações sobre Mortalidade (SIM) e o Sistema de Informação de Agravos de Notificação (SINAN), que apresenta informações complementares com a CAT, estando sobre administração do Ministério da Saúde, no qual coleta, processa e dissemina informações sobre saúde no Brasil por meio do DATASUS.

Existem diversas maneiras de obter informações sobre os acidentes de Saúde e Segurança do Trabalho (SST), sendo que duas abordagens principais na literatura analisada são o autorrelato e a análise retrospectiva dos dados (BOUCAUT; KNOBBEN, 2018). A análise dos dados históricos de incidentes é importante para identificar as principais causas, modos de falha, impactos relacionados e tendências estatísticas de tais acidentes, permitindo uma melhor compreensão dos mecanismos de incidentes e ajuda na preparação de medidas de prevenção e mitigação (GIRGIN; KRAUSMANN, 2016).

Embora tenha a obrigatoriedade do registro, quando os registros nos bancos de dados são analisados, a comunicação de acidente de trabalho nos órgãos regulatórios tem sido subnotificada conforme aponta vários estudos (HASSEL; ASBJØRNSLETT; HOLE, 2011, ABEGAZ *et al.*, 2014, SANGO *et al.*, 2016, SHORT; CAULFIELD, 2016, AHMED; SADULLAH; YAHYA, 2019, RIVERA; LAM, 2019, FELDMAN *et al.*, 2017).

No portal do SMARTLAB é destacado que no ano de 2018 ocorreram mais de 623 mil acidentes de trabalho e destes pelo menos 154 mil acidentes não foram emitidos a CAT, uma subnotificação de aproximadamente 25% dos acidentes na federação brasileira. Entre os motivos da subnotificação encontra-se a falta de conhecimento sobre como fazer o registro, o excesso de burocracia, o medo de punição pela chefia imediata e ao empirismo dos trabalhadores diante a necessidade de notificação devido à baixa gravidade do acidente (PROBST; GRASO, 2013, ZADOW *et al.*, 2017).

A prevalência de lesões no local de trabalho pode ocorrer quando (1) o trabalhador decide se denuncia a ocorrência da lesão ao empregador e (2) quando a empresa é notificada por um funcionário de uma lesão e decide se deve incluir essa lesão em seu registro oficial de lesão (PROBST; PETITTA; BARBARANELLI, 2017). A subnotificação no nível de organização envolve a última parte do fluxo e ocorre quando as organizações não relatam adequadamente as lesões relacionadas ao trabalho que atendem à definição de um incidente registrável à autoridade reguladora apropriada (PROBST; ESTRADA, 2010, PROBST; PETITTA; BARBARANELLI, 2017).

Correa; Assunção (2003) apresentam um estudo de subnotificação de acidentes de trabalhos fatais do município de Belo Horizonte, capital do estado de Minas Gerais, no qual os dados foram obtidos por meio do Sistema de Informação em Mortalidade (SIM), Sistema de Informação em Acidentes de Trabalho, do SUS de Belo Horizonte (SIAT-SUS/BH); e Sistema de Comunicação de Acidente de Trabalho, do Instituto Nacional de Seguro Social (INSS) referente ao ano de 1999. O estudo identificou a subnotificação por meio do cruzamento dos registros presentes em como nas três fontes de dados. Com dados de 2011, Bordoni *et al.*, (2016) mensuram a subnotificação de registros de acidentes fatais no Sistema de Informações sobre Mortalidade (SIM) novamente na cidade de Belo Horizonte, Minas Gerais. Neste estudo os autores utilizaram o método de captura e recaptura utilizando as fontes de dados do Sistema de Informação de Agravos de Notificação (Sinan) e Banco de Necropsias do Instituto Médico Legal de Belo Horizonte. Rodrigues; Santana (2019) estimam a subnotificação de acidentes de trabalho fatais nos sistemas de informação do Ministério da Saúde, Secretaria Estadual de Segurança Pública (SSP), Ministério do Trabalho e Secretaria da Previdência Social, de 2007 a 2015, na cidade de Palmas, capital do Tocantins. O estudo foi realizado por meio de pareamento dos casos de todos os sistemas de informação.

No presente trabalho a subnotificação dos registros será estimada por meio de técnica de captura-recaptura em todo o território nacional brasileiro, aumentando a abrangência em comparação com os estudos descritos anteriormente. Por meio da estimação da subnotificação busca particularmente responder a seguinte questão de pesquisa: “Qual é o índice de subnotificação de acidentes de trabalhos fatais quanto comparado os *datasets* da CAT, SINAN e SIM?”.

É concentrado a análise apenas nos dados de lesões fatais, uma vez que são tipicamente relatados com mais confiabilidade do que as lesões não fatais e são um ponto de partida para melhorar a vigilância de lesões ocupacionais (BENAVIDES *et al.*, 2003).

1.1 OBJETIVOS DO ESTUDO

O objetivo do presente trabalho está centrado na determinação do índice de subnotificação de acidentes de trabalhos fatais por meio da técnica de Captura e Recaptura nos bancos de registros de Saúde e Segurança do Trabalho brasileiros – CAT, SINAN e SIM.

1.1.1 Objetivos específicos

O objetivo da pesquisa se desdobra em alguns específicos, sendo:

- Realizar uma análise exploratória dos níveis atuais dos acidentes de trabalhos fatais entre os casos registrados e os subnotificados;
- Determinar as taxas de subnotificação de acidentes de trabalhos fatais entres os *datasets* estudados por meio do estimador de Lincoln-Petersen.

1.2 JUSTIFICATIVA

O sistema de saúde brasileiro passa a ter uma estrutura fragilizada com a subnotificação de acidentes e doença de trabalho, que é responsável por gerar prejuízos consideráveis ao passo que as doenças subnotificadas constituem risco à saúde da população.

Antes de determinar a subnotificação de acidentes de trabalho fatais é muito importante entender o comportamento das variáveis. As ferramentas de análise visual se tornaram muito populares para exploração de dados e obtenção de percepções a partir dos dados. Essas ferramentas são benéficas para explorar dados históricos de acidentes e analisar os fatores contribuintes para a ocorrência (ZHU *et al.*, 2021). Mahmood (2021) enfatiza que o processo de exploração dos dados é extremamente útil para realizar o entendimento básico do conjunto de dados, compreender a estrutura dos dados, as relações entre as variáveis dentro do conjunto de dados e as características gerais.

A subnotificação resulta em falta de evidências que possam ajudar a monitorar e revisar as medidas de SST no local de trabalho (PAPADOPOULO *et al.*, 2010). Sem estatísticas precisas sobre lesões no local de trabalho, os empregadores não podem identificar e começar a gerenciar adequadamente os problemas de saúde e segurança (AZAROFF; LEVENSTEIN; WEGMAN, 2002, TUCKER *et al.*, 2014). Por não entender a natureza das lesões, leva-se as organizações a ameaçarem a saúde e segurança a longo prazo dos seus trabalhadores (TUCKER *et al.*, 2014).

Para construir e desenvolver uma política pública, estatísticas precisas sobre lesões relacionadas ao trabalho são importantes para definir prêmios de seguros, comparar tendências de lesões entre grupos e jurisdições ocupacionais, desenvolver intervenções apropriadas de prevenção de lesões e, de maneira mais ampla, avaliar a necessidade de mudanças na legislação (SHANNON; LOWE, 2002, TUCKER *et al.*, 2014).

Tucker *et al.* (2014) enfatizam ainda sobre o comportamento na cultura social que a subnotificação de acidentes de trabalho pode causar. O autor condiciona a recusa em notificar acidentes de trabalho dos funcionários enquanto em idade jovem a uma continuidade de manter a prática na idade adulta e perpetuar as consequências negativas da subnotificação.

A preservação à integridade do trabalhador e a proteção a vida são os principais benefícios de determinar as taxas de subnotificação. Isto porque, altas taxas de não registo de acidentes fatais de acidentes de trabalho, gera nos órgãos de regulamentação a necessidade de rever a legislação presente e tomar ações para reduzir o índice de trabalhadores lesionados dentro das organizações.

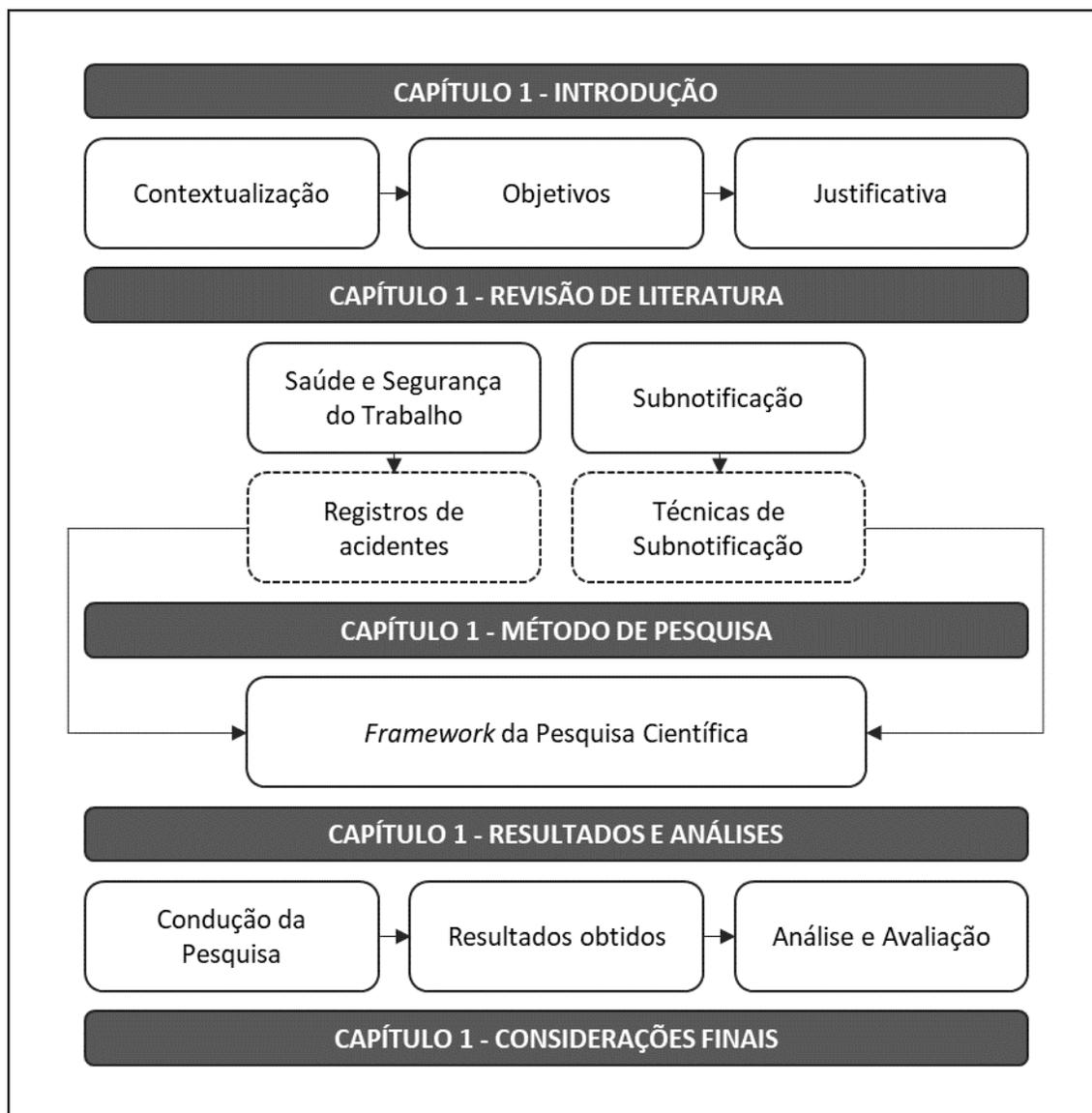
O estudo aborda os bancos de dados do SINAN e SIM, porém converge as análises para o banco de dados da CAT, devido ser o banco de dados de referência da federação brasileira para registros de acidentes de trabalho e ser base do INSS para o pagamento dos seguros aos envolvidos na ocorrência do incidente.

Responder à questão de pesquisa permitirá mapear o cenário mais próximo do real dos acidentes fatais dentro dos ambientes de trabalho na federação brasileira. A exploração da subnotificação primeiramente é fundamental para posicionar quanto a realidade dos casos não registrados nos órgãos oficiais. A análise das variáveis pode contribuir em avaliar os fatores que modelam este cenário e levantar ações reativas e preventivas que auxiliam tomadores de decisões a agir na fonte das causas.

1.3 ESTRUTURA DO TRABALHO

Após a introdução, o presente trabalho está estruturado em 4 capítulos, conforme Figura 1, que apresenta a relação de cada parte da pesquisa com as demais etapas do processo de construção da proposta do estudo.

Figura 1 – Visão geral da pesquisa



Fonte: Autor, 2021

No capítulo 2 é apresentado um referencial teórico sobre os temas centrais da pesquisa, que se encontra dividido em três seções. Primeiramente é introduzido uma revisão sobre Saúde e Segurança do Trabalho, no qual é apresentado uma abordagem geral sobre o tema e direcionando esta parte da revisão para os registros de acidentes de trabalho. A segunda seção aborda o assunto das subnotificações com enfoque no tema de Saúde e Segurança do Trabalho, trazendo definições sobre os assuntos e técnicas de aplicação.

A metodologia é apresentada no capítulo 3, que expõe a classificação da pesquisa e a descrição do método que moldou a condução do estudo. Posteriormente, no capítulo 4, é

apresentado um estudo exploratório dos bancos de dados em saúde e segurança do trabalho, realizando um comparativo entre algumas variáveis em comum, evidenciando a necessidade de analisar melhor o cenário de acidentes fatais no Brasil. Para finalizar a proposta de dissertação, é apresentado no capítulo 5 o cronograma de atividades futuras, descrevendo as próximas etapas do presente trabalho.

2

REVISÃO DE LITERATURA

Neste capítulo é apresentada a revisão bibliográfica sobre os temas centrais da pesquisa. A primeira parte da revisão é conceituado o tema de Saúde e Segurança do Trabalho dentro do escopo de acidente de trabalho e os registros dos acontecimentos nas bases brasileira. A segunda etapa aborda a definição de subnotificação com direcionamento voltado a bancos de dados de acidentes de trabalho.

2.1 SAÚDE E SEGURANÇA DO TRABALHO

A Saúde e Segurança do Trabalho compreende as condições e os fatores que afetam a saúde e a segurança dos trabalhadores e demais envolvidos no local de trabalho (DA SILVA; AMARAL, 2019, DEMIR; GUL; GUNERI, 2020). É papel da organização a garantia da segurança no local de trabalho e a disseminação de política de conscientização da importância de seguir as normas e procedimentos durante o trabalho às pessoas (SINGH; MISRA, 2020).

Diante das perspectivas da Saúde e Segurança do trabalho o conceito de competências de trabalho está baseado no equilíbrio entre as características individuais e as demandas de trabalho impostas ao trabalhador. Caso as características pessoais do trabalhador atendam às demandas do trabalho, o trabalhador poderá ter uma vida saudável. Contudo fatores físicos e psicossociais podem

comprometer a segurança e saúde do indivíduo caso não atendam às habilidades e expectativas do trabalhador (VARIANOU-MIKELLIDOU *et al.*, 2020).

Os sistemas de gerenciamento de saúde e segurança do trabalho são implementados nas empresas com o objetivo de permitir que as organizações forneçam ambientes de trabalho seguros e saudáveis, prevenindo lesões e problemas de saúde relacionados ao trabalho (ISO 45001, 2018). O novo padrão da ISO 45001 apresenta um foco maior no envolvimento do trabalhador, responsabilidade de gerenciamento e processos (MADSEN *et al.*, 2020). Quando o comprometimento e preocupação da alta administração com a segurança ocupacional são percebidos pela equipe os sistemas de gerenciamento em Saúde e Segurança do Trabalho podem apresentar uma maior eficácia (SKŁAD, 2019).

Para Ciarapica; Giancchetta (2009) é importante para as organizações abordar o gerenciamento de prevenção de forma sistemática e integrada às atividades da empresa por motivos como: (1) a exigência da legislação para este tipo de compromisso por parte da administração da empresa; (2) o requerimento de várias competências técnicas e funcionais que a prevenção exige; (3) o gerenciamento da prevenção para manter o nível de conflito dentro da norma e garantir que a organização esteja orientada para motivar a desenvolver recursos humanos e (4) o gerenciamento da segurança de forma correta para impedir a um aumento descontrolado dos custos, à deterioração dos relacionamentos com os sindicatos e à queda na imagem da empresa. Isso significa principalmente o papel que a gerência deve desempenhar para garantir a segurança no local de trabalho e aumentar a conscientização entre os trabalhadores sobre a importância de seguir as regras de segurança durante o trabalho (SINGH; MISRA, 2020).

A introdução dos conceitos de Saúde e Segurança do Trabalho afeta a redução de riscos e acidentes dentro das organizações, melhorando o gerenciamento de riscos e o cumprimento das obrigações legais, além de aumentar a produtividade, levando a um uso mais eficaz dos recursos e a uma maior capacidade de fornecer serviços e produtos consistente e ao mesmo tempo melhorando o desempenho financeiro (RAMOS; AFONSO; RODRIGUES, 2020, DEMIR; GUL; GUNERI, 2020; SINGH; MISRA, 2020). Apesar dos melhores esforços para minimizar o risco nas organizações, os acidentes parecem inevitáveis devido à “probabilidade de falha” inerente associada à tecnologia e ao sistema e à natureza inata das pessoas de cometer erros por omissão ou comissão (BEHERA; HASSAN, 2019).

Como a estimativa de mortalidade e morbidade relacionada ao trabalho em todo o mundo varia amplamente devido a problemas metodológico (TEUFER *et al.*, 2019), a percepção de risco é um preditor importante para desenvolver uma cultura de segurança adequada (RAMOS; AFONSO; RODRIGUES, 2020). Portanto, é essencial reduzir esses riscos examinando e analisando as causas dos acidentes e identificando riscos e perigos no local de trabalho (HAJAKBARI; MINAEI-BIDGOLI, 2014).

A conscientização dos trabalhadores leva mais em consideração os sentimentos do que conhecimentos ou avaliações objetivas dos riscos e os esforços de prevenção (RAMOS; AFONSO; RODRIGUES, 2020, ALBERT; PANDIT; PATIL, 2020), ou seja, os indivíduos tendem a superestimar os riscos com alta gravidade e subestimar os riscos com menor gravidade (RAMOS; AFONSO; RODRIGUES, 2020). A subestimação dos riscos de segurança aumenta probabilidade de exposição não intencional aos perigos na organização, que pode diminuir a sua produtividade (GUL; AK; GUNERI, 2017, TEUFER *et al.*, 2019, SINGH; MISRA, 2020) e conduzir a incidentes catastróficos (GUL; AK; GUNERI, 2017, ALBERT; PANDIT; PATIL, 2020). Os incidentes formados por doenças e acidentes de trabalho impactam a sociedade, devido as consequências que que os trabalhadores são submetidos, as empresas e a previdência social.

Por definição, acidentes são eventos aleatórios inesperados com resultados indesejáveis que levam à perda de produção, lesões ou problemas de saúde, perda de vidas, danos às propriedades e ao meio ambiente (BEHERA; HASSAN, 2019).

Os incisos 19 e 20 da Lei nº 8.213 de 1991 conceituam os termos acidente e doenças de trabalho ou ocupacional respectivamente. Acidente de trabalho ou de trajeto: é o acidente ocorrido no exercício da atividade profissional a serviço da empresa ou no deslocamento residência - trabalho - residência, e que provoque lesão corporal ou perturbação funcional que cause a perda ou redução – permanente ou temporária – da capacidade para o trabalho ou, em último caso, a morte. Doença ocupacional: é aquela produzida ou desencadeada pelo exercício do trabalho peculiar a determinada atividade e constante da respectiva relação elaborada pelo Ministério do Trabalho e da Previdência Social.

Por razão de irreversibilidade, pode-se considerar que a pior consequência do acontecimento de um acidente de trabalho ou o desenvolvimento de uma doença ocupacional é o óbito do trabalhador. Acidentes com graves consequências não são aceitáveis pela sociedade e precisam ser evitados por todos os meios possíveis (BEHERA; HASSAN, 2019). Acidentes e

doenças de trabalho resultam em um impacto extremamente significativo na vida dos trabalhadores.

2.1.1 Notificação de acidente de trabalho

O acidente de trabalho ou a doença ocupacional pode ser reportado aos órgãos oficiais tanto pela empresa quanto pelo trabalhador, vítima do acidente, ou médico que o atendeu. O conjunto destes registros é conhecido como banco de dados de acidentes de trabalho (AHMED; SADULLAH; YAHYA, 2019).

A Comunicação de Acidente de Trabalho – CAT – constitui o principal registro brasileiro para notificação de acidentes de trabalho e conforme abordado na introdução é obrigatório a abertura do comunicado junto ao INSS da notificação das doenças profissionais e das produzidas por condições especiais de trabalho, comprovadas ou suspeitas – Artigo 169 da CLT. A CAT pode ser realizada por qualquer pessoa, seja a empresa, quanto o funcionário, sendo necessário o preenchimento de quatro vias: a primeira do INSS, a segunda do segurado ou dependente, a terceira do sindicato de classe do trabalhador e a quarta via da empresa.

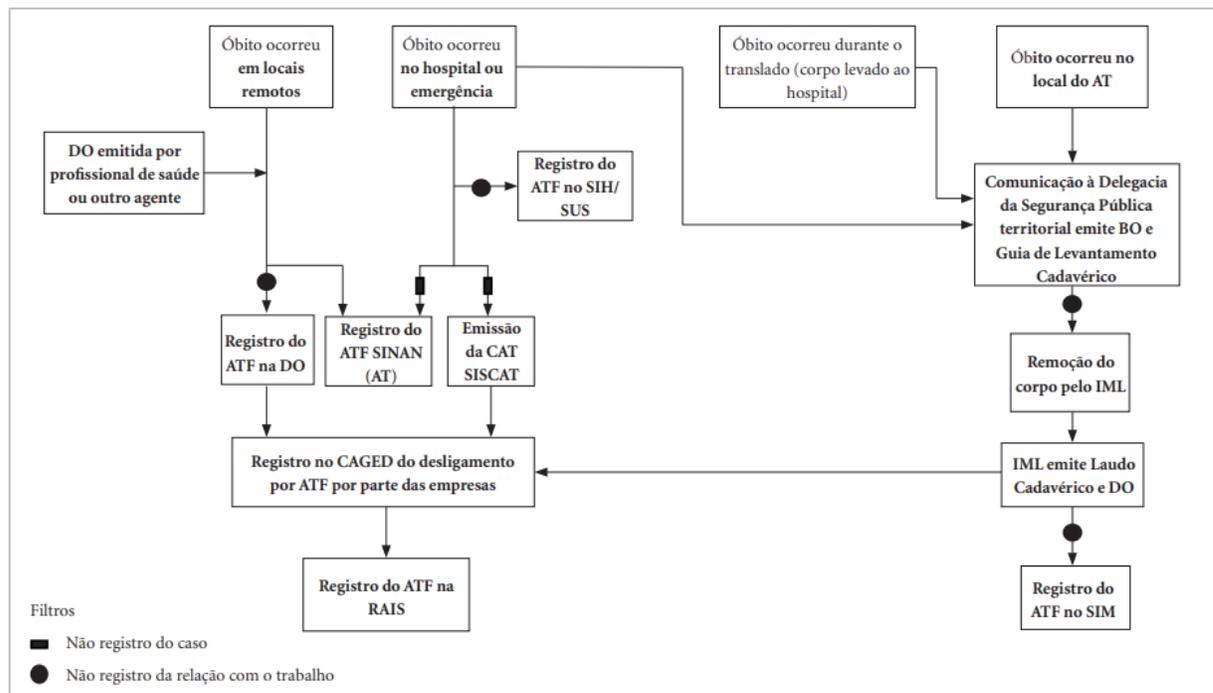
De acordo com o portal do INSS (2020) a emissão da CAT pode ser realizada em quatro momentos distintos:

- CAT inicial: refere-se ao acidente de trabalho típico, trajeto, doença profissional, do trabalho ou óbito imediato;
- CAT de reabertura: será utilizada para casos de afastamento por agravamento de lesão de acidente do trabalho ou de doença profissional ou do trabalho. Deverão constar as mesmas informações da época do acidente, exceto quanto ao afastamento, último dia trabalhado, atestado médico e data da emissão, que serão relativos à data da reabertura. Não será considerada CAT de reabertura a situação de simples assistência médica ou de afastamento com menos de 15 dias consecutivos.
- CAT de comunicação de óbito: será emitida exclusivamente para casos de falecimento decorrente de acidente ou doença profissional ou do trabalho, após o registro da CAT inicial.

2.1.2 Registros de acidentes de trabalho fatais

Batista; Santana; Ferrite (2019) apresentam a sistemática dos registros de Acidentes de Trabalhos Fatais (ATF) nas diferentes bases brasileiras. A Figura 2 apresenta o fluxograma de geração de dados nos bancos de dados relacionados a acidentes de trabalho.

Figura 2 - Fluxograma da geração de dados relacionados a Acidentes de Trabalhos Fatais



Legenda: AT: Acidente de Trabalho; ATF: Acidente de Trabalho Fatal; CAGED: Cadastro Geral de Empregados e Desempregados; CAT: Comunicação de Acidente de Trabalho; RAIS: Relação Anual de Informações Sociais; SISCAT: Sistema de Informação de Comunicações de Acidentes de Trabalho; BO: Boletim de Ocorrência; IML: Instituto de Medicina Legal; DO: Declaração de Óbito; SIH/SUS: Sistema de Informação Hospitalar do Sistema Único de Saúde; SIM: Sistema de Informação sobre Mortalidade; SINAN: Sistema de Informação de Agravos de Notificação.

Fonte: (BATISTA; SANTANA; FERRITE, 2019).

Os acidentes de trabalho fatais podem ocorrer em três circunstâncias: A primeira no local de ocorrência do acidente, no local de trabalho ou fora dele quando o funcionário está realizando uma tarefa de trabalho, no qual se faz necessária a comunicação imediata à delegacia de polícia, que registra a ocorrência e emite o boletim de ocorrência (BO) e a guia para o levantamento cadavérico, a ser realizado pela polícia técnica (BATISTA; SANTANA; FERRITE, 2019). O laudo cadavérico (LC) e a declaração de óbito (DO) são liberados após realização da necropsia e exames

laboratoriais quando necessário, que contém informações que podem ser essenciais para estabelecer a relação do acidente com o trabalho.

A segunda circunstância é quando o acidente ocorre durante o transporte ou quando em tratamento em serviço de emergência ou hospitalar – a instituição responsável e todos os passos descritos anteriormente devem ser seguidos, emissão do BO, LC e DO. Quando o atendimento médico foi fornecido pelo SUS, o ATF é registrado no Sistema de Informações Hospitalares (SIH/SUS) e em vários subsistemas do SINAN.

A terceira é quando a fatalidade acontece em locais remotos e distantes, onde não existem delegacia de polícia, IML ou serviço de saúde. Os procedimentos mudam de acordo com o contexto, mas em qualquer condição todos os documentos legais devem ser emitidos.

Os registros de acidentes de trabalhos fatais serão reunidos no SINAN caso houver unidades de atenção à saúde com equipe capacitadas onde os acidentes acontecem, independente da circunstância de ocorrência da morte. O INSS registra a ATF no SISCAT, quando a CAT é emitida e no SUIBE quando parentes recebem indenizações ou pensões, para qualquer caso de trabalhador formalmente registrado. O registro tanto no SISCAT quanto no SUIBE, está condicionado ao desejo e ao conhecimento do direito que a família tem de receber os benefícios e da equipe de saúde e dos empregadores.

De acordo com a Figura 2 no fluxo de registro de informações é apresentado alguns filtros que funcionam como barreiras que impedem que as informações sejam geradas conforme procedimento estabelecido, neste caso abre margem a possibilidade de gerar subnotificação de casos. Os autores apresentam duas estruturas de filtros, o primeiro a falta de registro da morte no sistema de informação e no segundo o tipo registra-se a morte, mas não a sua relação com o trabalho.

Batista, Santana e Ferrite (2019) reforçam que as causas desses filtros podem ser: (1) treinamento insuficiente de profissionais envolvidos; (2) pobre motivação e consciência sobre a relevância do registro da relação com o trabalho; (3) a preocupação com as implicações legais, especialmente entre aqueles responsáveis pelo registro do ATF; (4) a falta de equipamentos ou outros recursos necessários para notificar ou para investigar a relação do óbito com o trabalho e (5) as pressões de parte de empregadores, colegas, profissionais de saúde, advogados, e até mesmo familiares para a omissão do registro da relação com o trabalho devido a interesses pecuniários.

Conforme abordado, mesmo com a obrigatoriedade em registrar os acidentes e doenças ocupacionais nos órgãos regulatórios, diversas causas podem contribuir para a falta de registros nos *dataset* voltados a Saúde e Segurança do Trabalho, seja através na não notificação por parte da empresa quanto do funcionário. Conforme Boucaut e Knobben (2018) as informações sobre os incidentes de Saúde e Segurança do Trabalho pode ser obtida por meio da análise retrospectivas dos dados. Na presente pesquisa faz-se o uso desta abordagem para estudar a subnotificação dos bancos de dados relacionados a SSO.

2.2 SUBNOTIFICAÇÃO DE ACIDENTES DE TRABALHO

A subnotificação é definida em função do número de acidentes relatados pelo funcionário à organização e o número de acidentes sofridos pelo funcionário, mas não relatados à organização (PROBST; ESTRADA, 2010, PROBST *et al.*, 2018). A subnotificação não é adequadamente capturada simplesmente pelo número de acidentes não relatados, pois conforme aumenta a discrepância entre o número de acidentes não relatados e os relatados, a subnotificação aumenta (PROBST; ESTRADA, 2010).

O problema geral de subnotificar acidentes de trabalho e doenças ocupacionais se agrava em um ambiente de trabalho em mudança devido à pressão do tempo, insegurança no trabalho, presença de trabalhadores informais, imigrantes em local de trabalho e falta de vigilância médica para esses trabalhadores. (PAPADOPOULO *et al.*, 2010).

A subnotificação é explicada pela pressão do tempo na realização das atividades de trabalho e pelas dúvidas dos trabalhadores sobre elegibilidade, reputação, perda de renda e perspectivas de carreira. Embora cientes da subnotificação, os gerentes acreditam sutilmente nos comportamentos de risco moral dos trabalhadores, ou seja, que os trabalhadores não mudem seu comportamento de acordo com o contexto de acidente de trabalho (GALIZZI *et al.*, 2010). Embora grande maioria das organizações afirmam ter a Saúde e Segurança do Trabalho como uma preocupação primordial, as subnotificações de acidentes de trabalhos continuam ocorrendo com prevalência muito alta (PROBST; ESTRADA, 2010)

Entender a segurança do ambiente de trabalho requer análise das lesões não registradas e registradas com o objetivo de verificar as taxas reais de lesões e identificar os fatores que contribuem (ZADOW *et al.*, 2017). Fatores como gênero, setor econômico, tamanho da empresa,

a idade e a escolaridade do trabalhador, e as características de cada um, o trabalho, como longas horas de trabalho, monotonia, falta de autonomia no trabalho e insatisfação no trabalho são fatores importantes que incidem na taxa de acidentes relacionado ao trabalho (POULIAKAS; THEODOSSIOU, 2013).

A subnotificação de acidentes está relacionada ao clima de segurança organizacional, no qual um clima de segurança positivo pode reduzir os acidentes de trabalho (PROBST; ESTRADA, 2010). Embora algumas pesquisas tenham vinculado a subnotificação de acidentes a fatores como idade e posse, medo de represálias ou perda de benefícios, aceitação geral de que as lesões são um fato da vida em certas linhas de trabalho e pressão de produção (PROBST; GRASO, 2013).

O enfrentamento a COVID-19 tem mostrado que em países social e economicamente frágeis, a subnotificação coloca desafios adicionais a gestão da pandemia (REIS *et al.*, 2020), que também é considerada no Brasil como doença ocupacional. Assim, a limitação da estrutura federativa também pode ser considerado um fator para a taxa de subnotificação de acidentes ocupacionais.

O acesso a estatísticas precisas sobre lesões relacionadas ao trabalho é essencial para o desenvolvimento de programas eficazes de prevenção de lesões; no entanto, pesquisas envolvendo trabalhadores de todas as idades mostram que a subnotificação de lesões é predominante (TUCKER *et al.*, 2014). Um estudo de subnotificação requer um conjunto de parâmetros bem definidos do que constitui a subnotificação (HASSEL; ASBJØRNSLETT; HOLE, 2011).

Ahmed; Sadullah; Yahya (2019) enfatiza a extrema importância de conhecer os tipos de erros nos dados do acidente e os fatores que os causam, para que a solução correta possa ser identificada para o problema exato. O autor ressalta que erros de nível micro nos dados não pode ser solucionados com melhorias no nível macro, por exemplo: erros na fonte de onde os dados estão sendo gravado e transmitidos não podem ser solucionados com sistema avançado de gerenciamento de banco de dados. Contudo estes erros podem variar de país para país, devido a conjunto de fatores socioeconômico, demográfico e político, assim como os métodos para corrigir estes erros.

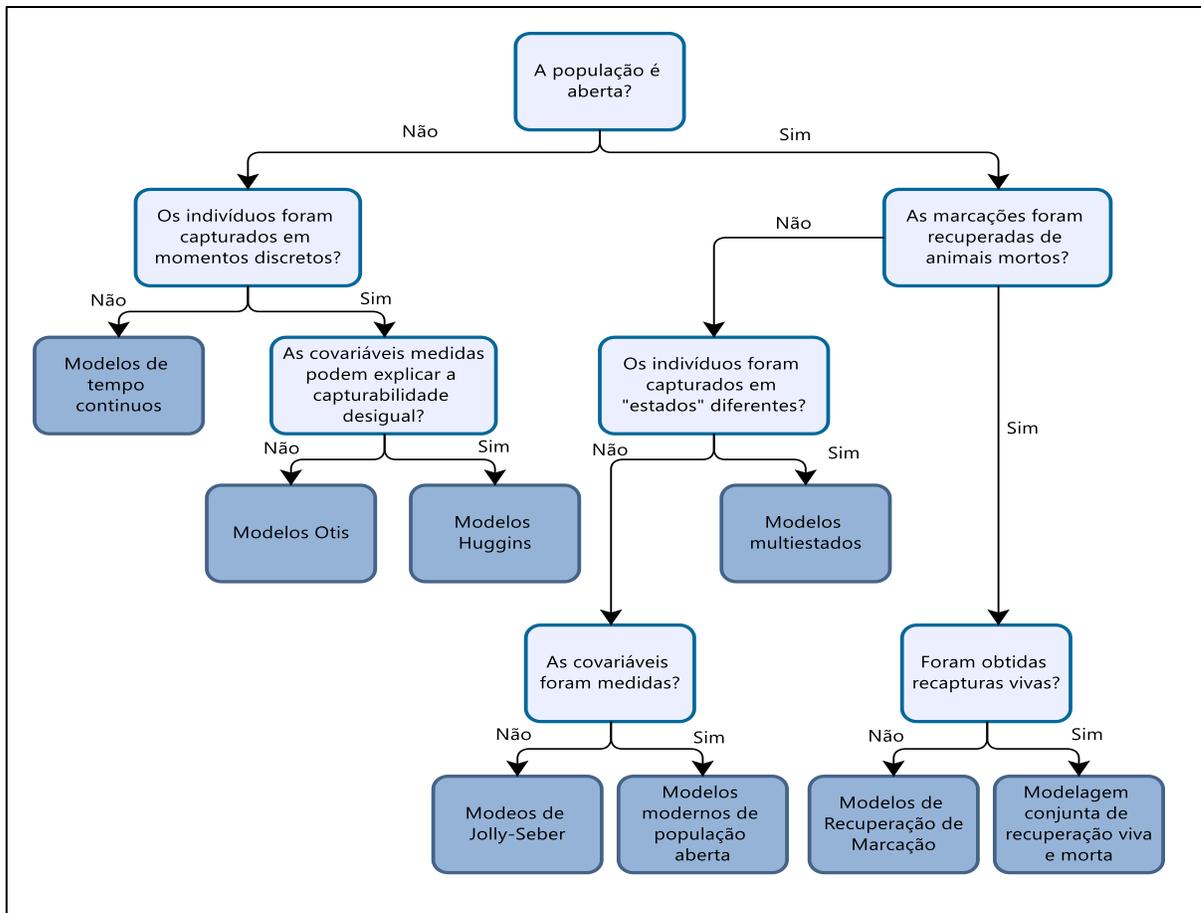
Em paralelo as recomendações passadas para os tomadores e mantenedores de registros a utilização de estimadores para corrigir os erros nos dados tem sido uma abordagem para aumentar a confiabilidade das bases de dados (AHMED; SADULLAH; YAHYA, 2019). Ahmed; Sadullah;

Yahya (2019) abordam os estimadores do método Captura e Recaptura para fazer inferência sobre uma amostra.

2.2.1 Captura e Recaptura

Em setembro de 1802, Pierre Simon Laplace fez-se uso de um tipo de abordagem de Captura e Recaptura para estimar o tamanho da população da França (AMSTRUP; MCDONALD; MANLY, 2010).

Figura 3 - Classificação dos modelos Captura e recaptura



Fonte: Amstrup, Mcdonald e Manly (2010)

O método de Captura e Recaptura pode ser utilizado para duas situações distintas: populações fechadas e populações abertas. É considerada população fechada quando o número total de indivíduos não muda por meio de nascimento, óbitos, imigração e emigração. Em ambiente real, na maioria de vezes a população não é fechada, porém quando a variação é pequena suficiente que

assumir como fechada é uma aproximação razoável, e os efeitos de violação assumidos são mínimos. Consequentemente população aberta é quando pode ser alterado durante o estudo, por qualquer combinação de nascimentos, óbitos, imigração ou emigração (AMSTRUP; MCDONALD; MANLY, 2010). A Figura 3 apresenta qual o momento e classificação do estimador adequado para cada situação.

A presente pesquisa focou-se em tratar de situações em que o método de captura e recaptura fosse empregado para população fechada, devido o comportamento das bases de dados e a temática que molda a pesquisa.

O método captura tem grande aplicação na epidemiologia para estimação da população de animais (BRITTAIN; BÖHNING, 2009, HASSEL; ASBJØRNSLETT; HOLE, 2011). Inicialmente o método marca indivíduos a partir de uma seleção de uma amostra e em seguida devolve ao seu ambiente. Em um segundo momento uma nova amostra é capturada, marcando os indivíduos marcados na primeira amostra. Caso a segunda amostra demonstre ser representativa da população total, a proporção de indivíduos marcados da segunda amostrada será aproximadamente igual a proporção dos indivíduos marcados na população total. A estimativa do número total de indivíduos na população é retirada da relação existente entre estas proporções (OLIVEIRA; 2007).

O método Captura e Recaptura tem sido amplamente utilizado para corrigir imprecisões nos relatórios de acidentes e estimar o número real de acidentes (ABEGAZ *et al.*, 2014, SANGO *et al.*, 2016, SHORT; CAULFIELD, 2016, AHMED; SADULLAH; YAHYA, 2019; RIVERA; LAM, 2019). O método permite realizar análise estratificada para avaliar o impacto de possíveis probabilidades de relatórios diferenciais de acordo com as variáveis envolvidas, como por exemplo: gravidade da lesão, sexo, idade (SANGO *et al.*, 2016).

Por meio do método de Captura e recaptura permitiu-se estimar o número total de óbitos nos bancos de dados governamentais (FELDMAN *et al.*, 2017), em mortes e lesões de trânsito (ABEGAZ *et al.*, 2014, SANGO *et al.*, 2016, SHORT; CAULFIELD, 2016, RIVERA; LAM, 2019), em acidentes marítimos (HASSEL; ASBJØRNSLETT; HOLE, 2011).

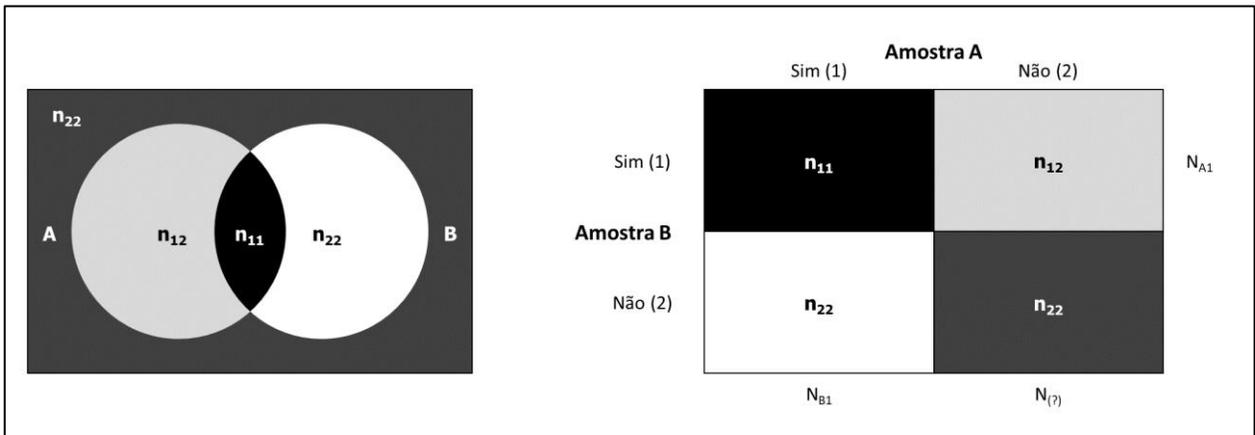
As quatro condições para o uso correto do método Captura-Recaptura são (HASSEL; ASBJØRNSLETT; HOLE, 2011, SHORT; CAULFIELD, 2016):

1. Que a população seja fechada;
2. Que haja uma identificação perfeita dos indivíduos;

3. Que as amostras sejam independentes e
4. Que em cada amostra a probabilidade de seleção seja constante.

O resultado de um processo de estimação do tamanho de uma população de duas fontes de dados pode ser obtido conforme Figura 4. Aplicação do método consiste na captura de uma amostra de indivíduos (amostra A) e em seguida captura uma nova amostra de indivíduos (amostra B) e identificado a quantidade de indivíduos recapturas nesta segunda amostragem.

Figura 4 - Diagrama de Venn e tabela de contingência de um processo de captura-recaptura envolvendo duas fontes



Fonte: Autor, 2021

Da Figura 4, tem:

$N = n_{11} + n_{12} + n_{21} + n_{22} =$ tamanho da população.

$N_{A1} = n_{11} + n_{12} =$ total de indivíduos capturados na primeira amostra (amostra A).

$N_{B1} = n_{11} + n_{21} =$ total de indivíduos capturados na segunda amostra (amostra B).

$n_{11} =$ número de indivíduos capturados na primeira e segunda amostras.

$n_{12} =$ número de indivíduos capturados somente na primeira amostra.

$n_{21} =$ número de indivíduos capturados somente na segunda amostra.

$n_{22} =$ número de indivíduos não capturados em nenhuma amostra.

O estimador de Lincoln-Petersen (LP) é frequentemente usado para estimar a população total de mortes ou feridos com base nos registros de duas fontes de dados. O tamanho da população é estimado utilizando o produto de dois conjuntos de dados (total de indivíduos ou elementos dos conjuntos) e dividindo pela sobreposição ou intersecção dos conjuntos (SHORT; CAULFIELD, 2016), dado pelo estimador de Lincoln-Petersen (N_{LP}), conforme Eq. (1):

$$N_{LP} = \frac{N_{A1} \times N_{B1}}{n_{11}} \quad \text{Eq. (1)}$$

Onde N_{A1} e N_{B1} são a quantidades de indivíduos capturados, respectivamente, na primeira e segunda amostras e n_{11} o número total de indivíduos capturados nas duas amostras. No entanto este estimador é viesado para amostras pequenas, levando a uma superestimativa do real tamanho populacional (SHORT; CAULFIELD, 2016). Adicionalmente, quando nenhum registro é correspondente no estimador de Lincoln-Petersen, esta abordagem falha e é necessária outra técnica de estimativa (HASSEL; ASBJØRNSLETT; HOLE, 2011).

A segunda alternativa é o estimador de Chapman que segundo Abegaz *et al.*, (2014) é menos afetado por zeros e é considerado menos tendencioso do que o estimador Lincoln-Peterson. O método Captura-recaptura, por meio do estimador de Chapman, por exemplo, possibilitou estimar o número de vítimas e o número de mortes em acidentes de trânsito (SANGO *et al.*, 2016, RIVERA; LAM, 2019) e a incidência corrigida de apuração de óbitos e lesões e a integridade das fontes de dados de tráfego e hospitalares (ABEGAZ *et al.*, 2014).

O estimador de Chapman (N_{CP}) é dado conforme a Eq. 2:

$$N_{CP} = \left[\frac{(N_{A1} + 1)(N_{B1} + 1)}{(n_{11} + 1)} \right] - 1 \quad \text{Eq. 2}$$

Onde N_{A1} e N_{B1} são as quantidades de pessoas capturadas na fonte A e B, e n_{11} o número de pessoas identificadas nas duas fontes. Caso as amostras não sejam independentes, N_{CP} será viesado. N_{CP} irá subestimar N_{LP} quando a probabilidade de captura na segunda amostra for maior do que na primeira, ou seja, uma dependência positiva. N_{CP} irá superestimar N_{LP} quando a probabilidade de captura na segunda amostra for menor do que na primeira, ou seja, uma dependência negativa (COELI; VERAS; COUTINHO, 2000).

Um dos grandes desafios de trabalhar com diferentes bancos de dados é a padronização dos dados, que leva a necessidade de trabalhar com técnicas mais avançadas para transformar os dados. Dado o fundamento do método de Captura-Recaptura em ter a identificação exata entre os bancos de dados, trabalhar no tratamento dos dados pode auxiliar neste processo em classificar os registros quando não tem uma variável que correlaciona os diferentes *dataset*. No tópico a seguir é apresentado uma fundamentação teórica sobre seleção e pré-processamento de dados.

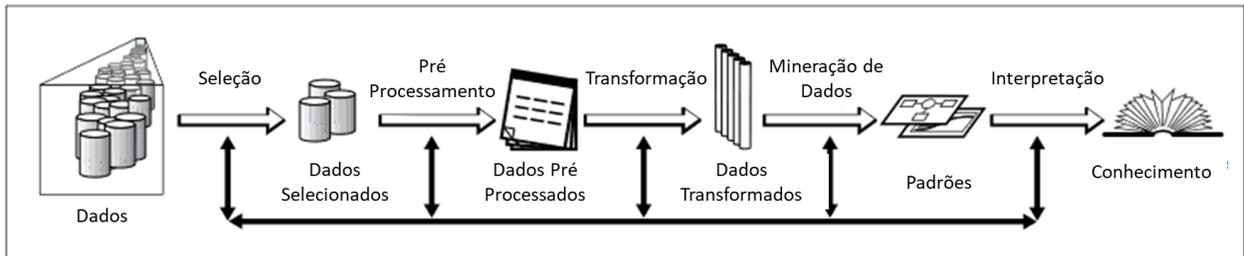
2.3 SELEÇÃO E PRÉ-PROCESSAMENTO DE DADOS

Considerando o enorme capital humano e as perdas financeiras decorrentes de lesões, os pesquisadores continuamente buscam maneiras de compreender melhor os fatores que afetam a ocorrência e a gravidade dos incidentes (KAKHKI; FREEMAN; MOSHER, 2019). Em particular, antes de iniciar uma análise complexa da situação no local de trabalho, seria melhor entender completamente a estrutura de dados para identificar os relacionamentos intrínsecos entre as diferentes variáveis envolvidas e, assim, extrair regras úteis de tomada de decisão (BEVILACQUA; CIARAPICA; GIACCHETTA, 2008).

O uso da tecnologia de dados tem sido extensivamente aplicado em diversos domínios (LIAO; PERNG, 2008, AIT-MLOUK; GHARNATI; AGOUTI, 2017, KODAPANAKKAL *et al.*, 2020), incluindo acidentes de trabalho, que com o crescente uso de bancos de dados, passam da função de auxílio de processamento de dados às principais armas estratégicas para prevenção de lesões (LIAO; PERNG, 2008). O desenvolvimento de novas Tecnologia de dados ou *Data Technology*, é decorrente da necessidade da facilidade de processamento, armazenamento e compartilhamento de grandes volumes de dados (KODAPANAKKAL *et al.*, 2020). O grande volume e a alta dimensionalidade dos bancos de dados de acidentes levam a um colapso na análise humana tradicional, tornando difícil o entendimento, considerando o número de variáveis envolvidas (BEVILACQUA; CIARAPICA; GIACCHETTA, 2008, LIAO; PERNG, 2008).

Trabalhar com um volume expressivo de dados é necessário um direcionamento para conduzir no processo da coleta de dados até a entrega de conhecimento extraído das, e uma das metodologias mais utilizadas é o Processo de Descobrimto de Conhecimento, mais comumente conhecido como KDD (*Knowledge Discovery in Databases*). De acordo com Fayyad (1996) o KDD é um processo de várias etapas, não trivial, interativo e iterativo, para identificação de padrões compreensíveis, válidos, novos e potencialmente úteis a partir de grande conjunto de dados. O processo KDD está representado na Figura 5, que apresenta a condução interativa e iterativa do processo.

Figura 5 - Visão geral do Processo de Descobrimto de Conhecimento



Fonte: Adaptado, Fayyad (1996)

A etapa de seleção do processo KDD inclui a seleção de um conjunto de dados ou o foco em um subconjunto de variáveis ou amostras de dados nos quais a descoberta deve ser realizada. A etapa de processamento constitui a inclusão de operações básicas, como remover ruídos ou outliers e decidir estratégias para lidar com campos de dados ausentes. Na etapa de transformação visa encontrar recursos úteis para representar os dados e usar métodos de redução de dimensionalidade ou transformação para reduzir o número efetivo de variáveis em consideração para encontrar representações invariantes para os dados. A etapa de Mineração de Dados inclui a seleção da técnica a ser utilizada, a escolha do algoritmo, e pôr fim a execução da mineração de dados que constituem a pesquisa por padrões de interesse em uma forma representacional específica ou em um conjunto de tais representações com base nas técnicas de mineração. Por fim é realizada a interpretação dos padrões descobertos e, possivelmente, o retorno a qualquer uma das etapas anteriores, bem como a possível visualização dos padrões extraídos, a remoção de padrões redundantes ou irrelevantes e a tradução dos termos úteis compreensíveis pelos usuários (FAYYAD, 1996).

2.4 TRABALHOS CORRELATOS

Neste tópico é apresentado estudos correlatos de subnotificação dentro do contexto de acidentes de trabalho. Realizou-se uma pesquisa de trabalhos que apresentassem ambos os temas em diferentes áreas de atuação.

Na área da saúde a subnotificação de acidentes de trabalho é estudada por meio de questionários, dentro do corpo de profissionais que estão expostos a materiais biológicos ou mediante contato com mucosas (NAGAO *et al.*, 2009, VOIDE *et al.*, 2012, DIAS FERREIRA *et al.*, 2015, BARBOSA *et al.*, 2017).

Por meio de entrevista dos detentores de registros de lesões e doenças ocupacionais nos estabelecimentos do Estado de Washington, foram aplicados métodos qualitativos e quantitativos para explorar a manutenção de registros e práticas de negócios que podem explicar a captura incompleta de casos entre dois bancos de dados nacionais (WUELLNER; BONAUTO, 2014). A subnotificação de lesões ou doenças relacionadas ao trabalho à remuneração dos trabalhadores foi quantificada por meio das técnicas de regressão logística utilizando dados de uma entrevista com trabalhadores que receberam salário e participaram do Sistema de Vigilância de Fatores de Risco Comportamental do Estado de Washington em 2002 (FAN *et al.*, 2006).

Dentro da faixa dos 25 anos de idades de jovens canadenses foi explorada a subnotificação de acidentes de trabalho com ritmo de trabalho de meia jornada, com o objetivo de fornecer uma estimativa da prevalência de acidentes de trabalho com afastamento nessa. Dez razões foram fornecidas pelos entrevistados para evitar o relato, com as mais comuns sendo percebidas baixa gravidade da lesão, reações negativas de outras pessoas, como empregadores, e ambiguidade sobre se a lesão foi causada pelo trabalho (TUCKER *et al.*, 2014).

A subnotificação de lesões ocupacionais foi examinada em quatro unidades de saúde norte americana, usando dados quantitativos, qualitativos e observacionais, testando empiricamente tanto a magnitude da subnotificação quanto os fatores relacionados a lesões ocupacionais e subnotificação (GALIZZI *et al.*, 2010).

Dentro do contexto organizacional é examinado os preditores de subnotificação de acidentes de funcionários em vários setores da indústria e testado até que ponto o clima de segurança organizacional e as percepções da fiscalização das políticas de segurança pelo supervisor estão relacionadas a uma falha no relato de acidentes (PROBST, 2015).

No contexto de acidentes de trabalhos fatais no Brasil, alguns estudos foram realizados. Correa; Assunção (2003) apresentam um estudo de subnotificação de acidentes de trabalhos fatais do município de Belo Horizonte, capital do estado de Minas Gerais, no qual os dados foram os obtidos por meio do Sistema de Informação em Mortalidade (SIM), Sistema de Informação em Acidentes de Trabalho, do SUS de Belo Horizonte (SIAT-SUS/BH); e Sistema de Comunicação de Acidente de Trabalho, do Instituto Nacional de Seguro Social (INSS) referente ao ano de 1999. O estudo identificou a subnotificação por meio do cruzamento dos registros presentes em como nas três fontes de dados. Com dados de 2011, Bordoni *et al.*, (2016) mensuram a subnotificação de registros de acidentes fatais no Sistema de Informações sobre Mortalidade (SIM) novamente na

cidade de Belo Horizonte, Minas Gerais. Neste estudo os autores utilizaram o método de captura e recaptura utilizando as fontes de dados do Sistema de Informação de Agravos de Notificação (Sinan) e Banco de Necropsias do Instituto Médico Legal de Belo Horizonte. Rodrigues; Santana (2019) estimam a subnotificação de acidentes de trabalho fatais nos sistemas de informação do Ministério da Saúde, Secretaria Estadual de Segurança Pública (SSP), Ministério do Trabalho e Secretaria da Previdência Social, de 2007 a 2015, na cidade de Palmas, capital do Tocantins. O estudo foi realizado por meio de pareamento dos casos de todos os sistemas de informação.

Pode-se afirmar que o estudo de subnotificação tem sido explorado ao longo dos anos e o presente trabalho tem como finalidade compor este portfólio de estudos, centralizando os esforços em definir os acidentes e doenças subnotificados e apresentar os fatores que contribuem para estes registros.

MÉTODO DE PESQUISA

Neste capítulo é abordado o método de pesquisa que caracteriza a pesquisa e detalha as etapas do método que conduz o desenvolvimento da pesquisa.

3.1 CARACTERIZAÇÃO DA PESQUISA

A Pesquisa Científica tem como finalidade responder perguntas e conhecer cientificamente um ou mais aspectos de determinado assunto. A pesquisa científica deve ser sistemática, metódica e crítica e o produto da pesquisa científica deve contribuir para o avanço do conhecimento humano (MARCZYK; DEMATTEO, 2005, PRODANOV; FREITAS, 2013).

A pesquisa pode ser classificada pela natureza, abordagem, objetivos e procedimentos (GIL, 2002; GERHARDT; SILVEIRA, 2009, PRODANOV; FREITAS, 2013), sendo utilizada esta estrutura para classificar o presente estudo. Com natureza aplicada a pesquisa tem como objetivo gerar conhecimentos para aplicação prática, dirigidos à solução de problemas específicos, envolvendo verdades e interesses locais.

Quanto aos objetivos o presente estudo pode ser classificado em três categorias. A fase preliminar da pesquisa é classificada como exploratória, tendo como finalidade proporcionar mais informações sobre o assunto que pretende investigar e facilitar a delimitação do tema das pesquisas. Para atender este objetivo, o levantamento bibliográfico é utilizado para contextualizar a pesquisa

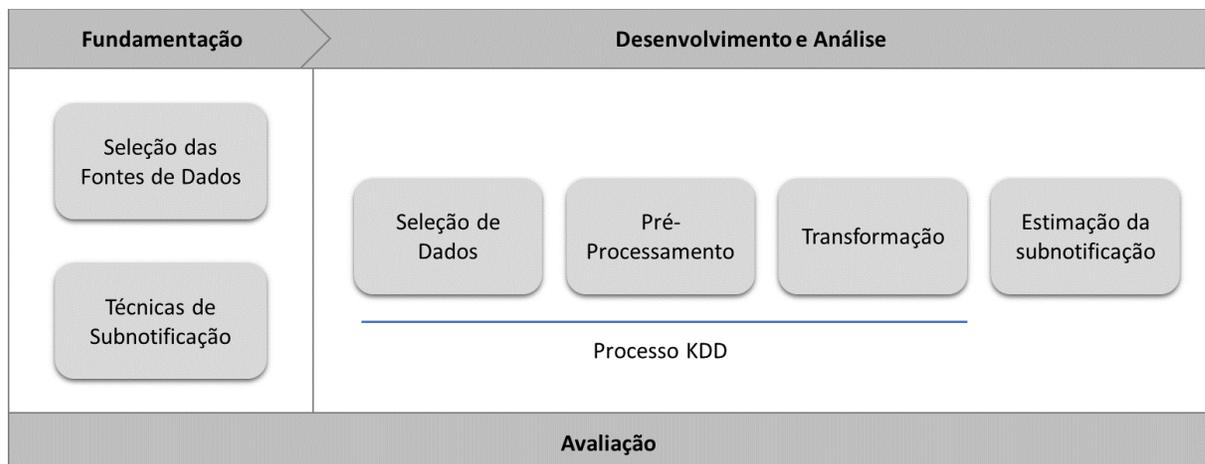
e determinar os métodos de subnotificação de casos. No segundo momento passa para a fase descritiva que tem como objetivo descrever os fatos e fenômenos que determinam a realidade. Finalizando com a pesquisa explicativa, sendo sequência da descritiva, que visa identificar os fatores que determinam ou que contribuem para a ocorrência dos fenômenos.

Quanto aos procedimentos a pesquisa é classificado inicialmente como pesquisa bibliográfica por realizar o levantamento de referências teóricas já analisadas, permitindo ao pesquisador conhecer o que já se estudou sobre o assunto com o objetivo de recolher informações ou conhecimentos prévios sobre o problema a respeito do qual se procura a resposta. posteriormente a pesquisa passa a ser classificada como *Ex-Post-Facto* devido o objetivo de investigar possíveis relações de causa e efeito entre um determinado fato e um fenômeno que ocorre posteriormente, tendo como principal característica o fato de os dados serem coletados após a ocorrência dos eventos.

3.2 MÉTODO DE ESTUDO

A metodologia do estudo foi construída a partir da exploração da temática da Saúde e Segurança do Trabalho e a Subnotificação voltada a acidentes e doenças de trabalho, conforme Figura 6.

Figura 6 - Metodologia de Pesquisa



Fonte: Autor, 2021

Estes assuntos que delimitam a pesquisa foram fundamentais para selecionar as melhores técnicas de subnotificação e preparar os bancos de dados para sua aplicação. As técnicas foram

identificadas por meio da análise de trabalhos que abordam a subnotificação em bancos de dados com mesma estrutura do presente estudo, porém não necessariamente com aplicação em Saúde e Segurança do Trabalho. E os bancos de dados foram selecionados por meio da exploração e entendimento da dinâmica de como ocorrem os registros e identificar os pontos fracos.

A revisão literária permitiu entender melhor o contexto de como ocorre os registros de acidentes de trabalho nas diversas bases de dados de acordo com a finalidade e propósito de cada uma. A avaliação de cada uma destas bases permitiu determinar quais as bases poderiam atender os objetivos da pesquisa. Com base nas descrições nos portais que os registros encontram hospedados, selecionou-se as bases para compor um banco de registros. Para o presente estudo os bancos de dados da CAT, SINAN e SIM foram consideradas para análise da subnotificação de acidentes fatais em local de trabalho durante o ano de 2019.

O segundo enfoque da revisão de pesquisa referiu-se ao tema de subnotificação de acidentes fatais e não fatais em locais de trabalho. Os resultados encontrados apresentaram técnicas que podem ser utilizadas para estimar o nível de subnotificação de acidentes no sítio de estudo. Catalogou-se as técnicas encontradas para auxiliar no momento de aplicação nos bancos de dados encontrados. A proposta de dissertação apresenta uma das principais técnicas para determinar a subnotificação de casos, o método Captura-Recaptura, com a aplicação de dois estimadores distintos, que serão considerados nesta pesquisa. O estudo deve continuar aprofundando o levantamento de técnicas para determinar as que melhor contribuem com a relevância do presente trabalho.

Com o levantamento das fontes de dados e das técnicas é iniciado o terceiro enfoque da pesquisa que caracteriza o processo análise dos dados para determinar a subnotificação de mortes classificadas dentro do contexto de segurança de trabalho. As etapas do processo de análise de dados são baseadas nas etapas do *Knowledge Discovery Data (KDD)*.

O Modelo proposto é interativo, devido a combinação de homem - máquina e iterativo, devido a refinamento sucessivos e subsequentes. Isto é, independe da etapa que a pesquisa estiver sendo executada, pode-se voltar para etapas anteriores com o objetivo de ajustar diretrizes para finalização da obtenção das respostas. A seguir é detalhado cada etapa do desta etapa da pesquisa.

Com a o portfólio de banco de dados formado, a primeira etapa do processo é a seleção de dados que tem como objetivo a redução do banco de dados para um tamanho ideal de trabalho,

identificando quais informações, dentre as bases de dados existentes devem ser consideradas durante o processo KDD.

O pré-processamento consiste em atividades que visam preparar os dados para os algoritmos executarem os comandos. Nesta etapa dados ausentes, errôneas ou inconsistentes nas bases são corrigidas para garantir a qualidade das informações geradas.

A etapa de transformação consiste na transformação dos dados de forma que o algoritmo possa interpretar na ferramenta a ser utilizada. Os bancos de dados não apresentam a mesma classificação para as variáveis, por exemplo, no banco de dados do SIM a idade vem codificada em uma expressão alfanumérica e não em número decimal. Outro exemplo é a variável sexo que em alguns bancos de dados está codificada em forma textual e outros em forma numérica. Assim é necessário decodificar algumas variáveis para padronizar com as demais.

Como abordado no referencial teórico, a técnica captura e recaptura, tem como uma das premissas a identificação mais assertiva possível dos dados que estão sendo analisados. No Brasil o documento que garante a identificação de uma pessoa é Cadastro de Pessoa Física (CPF), contudo este número de documento não está disponível nos bancos de dados, nem como nome do acidentado. Para determinar a identificação dos trabalhadores que sofreram a fatalidade realizou a concatenação das variáveis com o objetivo de criar um código rastreável entre as fontes de dados.

Com a base preparada a próxima etapa é a determinação da subnotificação por meio do método de captura e recaptura, que apresenta algumas premissas importantes para a aplicação:

1. Que a população seja fechada;
2. Que haja uma identificação perfeita dos indivíduos;
3. Que as amostras sejam independentes e
4. Que em cada amostra a probabilidade de seleção seja constante.

Durante a exploração dos dados para análise descritiva dos dados, inicialmente fez-se uso da ferramenta Excel[®] da Microsoft para compilação e transformação dos dados das diferentes fontes. A ferramenta Power BI[®], Microsoft, foi utilizada para criar relacionamento com base de dados auxiliares, como de cidades do IBGE e da tabela geral de CNAE, além da elaboração de gráficos informativos.

A última etapa da pesquisa consiste na avaliação dos resultados obtidos e levantamento de considerações importantes obtidas com a aplicação do método. O tópico seguinte apresenta uma análise exploratória dos dados..

4

ANÁLISE EXPLORATÓRIA

O estudo tem a finalidade de analisar a subnotificação dos registros nos bancos de dados de acidentes e doenças ocupacionais fatais. Porém, conforme abordado na justificativa do trabalho, tem-se um enfoque para os registros do banco de dados da CAT. Neste capítulo é apresentada uma análise exploratória a partir da coleta e pré-processamento de registros dos bancos de dados de acidentes fatais de trabalho.

4.1 DADOS DA SAÚDE E SEGURANÇA DO TRABALHO

Conforme artigo 8º da Lei 12.527, de 18 de novembro de 2011, Lei de Acesso à Informação (LAI), é dever dos órgãos e entidades públicas promover, independentemente de requerimentos, a divulgação em local de fácil acesso, no âmbito de suas competências, de informações de interesse coletivo ou geral por eles produzidas ou custodiadas. Dentro deste amparo legal é possível ter acesso aos principais bancos de dados que registram as letalidades decorrentes de acidentes ou doenças de trabalho. Os bancos de dados do CAT, SINAN e SIM constituem os principais *dataset* de registros de acidentes fatais, e são os que foram considerados para o presente estudo.

O DATAPREV é uma empresa pública brasileira da previdência, vinculado ao Ministério da Economia, responsável por manter a atualização dos registros de Comunicação de Acidente de

Trabalho (CAT). De acordo com o portal do INSS (2020) a CAT é um documento emitido para reconhecer tanto um acidente de trabalho ou de trajeto assim como uma doença ocupacional.

O DATASUS é o Departamento de Informática do Sistema Único de da Saúde (SUS), vinculado ao Ministério da Saúde que tem como responsabilidade por prover os órgãos do SUS de sistemas de informação e suporte de informática, necessários ao processo de planejamento, operação e controle. Entre os bancos de dados que são disponibilizados pelo DATASUS, estão os do SINAN e SIM que apresentam registros relacionados a Acidente de Trabalho.

De acordo com a portal eletrônico do DATASUS (2020), o banco de dados do SIM foi criado pelo Ministério da Saúde com o objetivo de obtenção regular de dados sobre mortalidade no país. A partir da criação do SIM foi possível a captação de dados sobre mortalidade, de forma abrangente, para subsidiar as diversas esferas de gestão na saúde pública. Com base nessas informações é possível realizar análises de situação, planejamento e avaliação das ações e programas na área.

Segundo o portal eletrônico do SINAN (2020) o Sistema de Informação de Agravos de Notificação (SINAN) é alimentado, principalmente, pela notificação e investigação de casos de doenças e agravos que constam da lista nacional de doenças de notificação compulsória (Portaria de Consolidação nº 4, de 28 de setembro de 2017, anexo V - Capítulo I), mas é facultado a estados e municípios incluir outros problemas de saúde importantes em sua região.

Embora cada banco de dados tenha um enfoque específico, todos eles registram dados referentes a casos mais graves de acidentes relacionados ao trabalho. O banco de dados da CAT, tem um foco na área de acidentes de trabalho, portanto é possível caracterizar melhor cada caso com as variáveis deste *dataset*. Os bancos de dados do SUS, principalmente o SIM, tem um foco voltado mais para área hospitalar, sendo assim, as limitações quanto a caracterização do acidente de trabalho são altas. A Tabela 1 apresenta uma descrição de como está constituído cada banco de dados e quais são os grupos de campos presentes.

Tabela 1 - Caracterização dos bancos de dados de SSO no Brasil

Grupos de Campos	CAT	SIM	SINAN
Dados Gerais	✓	✓	✓
Dados Pessoais	✓	✓	✓
Dados Residenciais	✓	✓	✓
Dados da empresa	✓	✗	✓
Dados Básico do Acidente	✓	✓	✓
Dados Complementares do Acidente	✓	✗	✓
Dados do Atendimento Médico	✗	✓	✓
Conclusão do Caso	✗	✓	✓

Fonte: Autor, 2021.

A construção dos Grupos de Campos ocorreu por meio do levantamento de todas as variáveis dos *datasets* selecionados para análise, e agrupando conforme a característica de cada uma. Posteriormente realizou a classificação de qual categoria de variável estava presente em cada *dataset*. A descrição de cada grupo de campo é dada por:

- Dados Gerais: correspondem a notificação, quando foi aberta;
- Dados Pessoais: os campos referentes a caracterização da pessoa, como sexo, idade, data de nascimento e demais informações quando presente, como endereço, raça, entre outros;
- Dados Residenciais: refere-se ao endereço do acidentado;
- Dados da Empresa: descrição da empresa quando presente, como CNAE, segmento, CNPJ e demais informações;
- Dados Básico do Acidente: Refere-se à apresentação da categorização do registro como acidente de trabalho;
- Dados Complementares do Acidente: quando apresenta a caracterização das variáveis do acidente ou doença de trabalho;
- Dados do Atendimento Médico: Informações referente ao atendimento do incidente, como internação por exemplo;
- Conclusão do caso: quando o banco de dados apresenta a evolução do caso.

Mesmo alguns bancos não tendo informações importantes referente ao acidente, todos eles são relevantes para estimar a real situação de acidentes de trabalhos fatais no Brasil. A explanação do processo de registro de cada evento nos bancos de dados é uma forma de compreender a

finalidade de cada *dataset* e avaliar e identificar as oportunidades de estudo para gerar conhecimento aos tomadores de decisão.

4.2 PREPARAÇÃO DOS DADOS

O estudo exploratório analisa os registros das principais características de dois bancos de dados além da CAT, sendo: SINAN e SIM. O período de busca foi concentrado no ano de 2019, devido ser o único período que coincide com a disponibilização dos dados em todas as fontes de dados. Os registros da CAT iniciam no segundo semestre de 2018, e a atualização mais recente que coincide as três fontes de dados caracterizou o ano de 2019.

O banco de dados da CAT apresenta registros de acidentes e mortes de trabalho, já o do SINAN apresenta os casos graves e óbitos e o do SIM apresenta apenas óbitos relacionado a acidente de trabalho. A primeira etapa foi preparar e padronizar os dados e manter em todos eles apenas os óbitos.

Após a seleção dos dados, agrupou-se todos os arquivos em um único. Como os bancos apresentam notações próprias, realizou-se a padronização dos atributos de cada banco para um geral. A Tabela 2 apresenta a padronização para as variáveis utilizadas na pesquisa.

Tabela 2 - Padronização das variáveis das fontes de dados

Variáveis	CAT	SINAN	SIM	Padrão
Datas	dd/mm/aaaa	dd/mm/aaaa	ddmmaaaa	dd/mm/aaaa
Sexo	Masculino	M (Masculino)	1 – Feminino	M
	Feminino	F (Feminino)	2 – Masculino	F

Fonte: Autor, 2021.

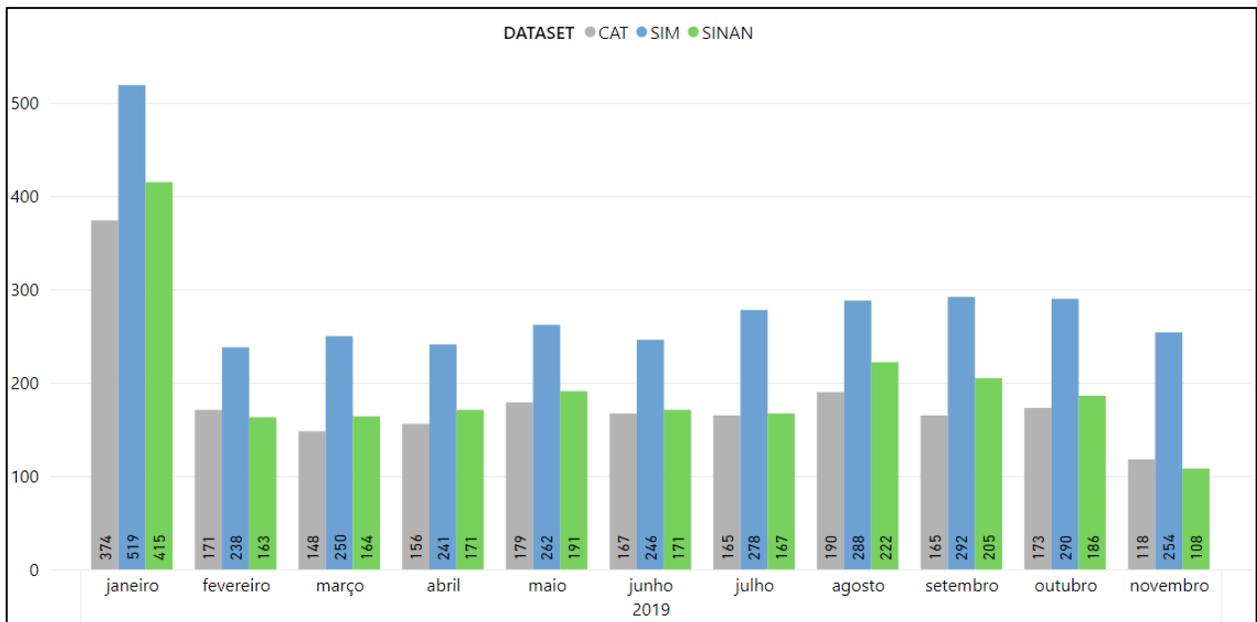
Da mesma forma ocorreu para variáveis de datas e idades que precisaram serem decodificadas, ou o enriquecimento de variáveis com informações incompletas, como por exemplo o CID-10 das classificações de doenças e o CNAE das classificações das empresas.

Na seleção dos dados, retirou os registros que não apresentavam a data de nascimento ou o sexo. Como em alguns *datasets* apresentam apenas o sexo masculino ou feminino como padrão de registro, retirou dos demais quando o sexo da pessoa não está definido entre feminino ou masculino.

4.3 ANÁLISE EXPLORATÓRIA DOS DADOS

Com os dados devidamente selecionados, os bancos de dados da CAT, SINAN e SIM resultaram em 2.006, 2.163 e 3.159 registros respectivamente. Conforme a Figura 7 percebe-se que o comportamento de crescimento e decrescimento ao longo dos seis meses são similares entre os bancos de dados, contudo a variabilidade na quantidade de registros entre os bancos é consideravelmente expressiva. O alto índice no primeiro mês do ano é justificado por um dos maiores acidentes ambientais e humano da história brasileira através do rompimento da barragem de extração mineral de Brumadinho, uma cidade do interior do estado de Minas Gerais, que resultou num alto número de vítimas além do impacto causado na região.

Figura 7 – Quantidade de registros por base de dados



Fonte: Autor, 2021

O banco de dados da CAT, que é o foco deste trabalho, apresenta a menor quantidade de registros de acidentes de trabalho entre os *datasets*. Os registros do SINAN são 7,8% maior que o da CAT e os registros do SIM apresenta um percentual de 57,5% maior que o da CAT. Importante ressaltar que este número evidencia o valor mínimo da subnotificação, pois a quantidade de registros que os bancos de dados têm iguais entre sim, referem-se ao mesmo registro.

Ressalta-se que o banco de dados do SIM apresenta o maior número de caso em comparação com os demais bancos, especialmente em relação a CAT e SINAN que possui foco nos registros

de acidentes de trabalho, enquanto o banco de dados do SIM tem como foco os registros em que os funcionários precisaram de atendimento hospitalares. Contudo, mesmo a base apresentando os maiores números de registros, não se pode inferir que todos os dados das demais estão presente na mais completa. Dentro da base do SIM pode ter registros faltantes que estão presente nas demais. Assim, realizar um estudo com o maior número de bases aumenta a acuracidade da determinação as taxas de subnotificação de uma base em comparação com as outras.

Estratificando alguns fatores dos bancos de dados, a idade das pessoas demonstra um importante agravante na ocorrência de letalidade nos locais de trabalho, e quando aplicado a variável sexo é nítido o comportamento da incidência entre os trabalhadores masculino e femininos. A Tabela 3 apresenta a quantidade de acidentes fatais por sexo para cada fonte de dados, podendo notar embora apresentam uma quantidade significativa de uma base para outra, a representatividade do sexo dentro de cada base é praticamente similar.

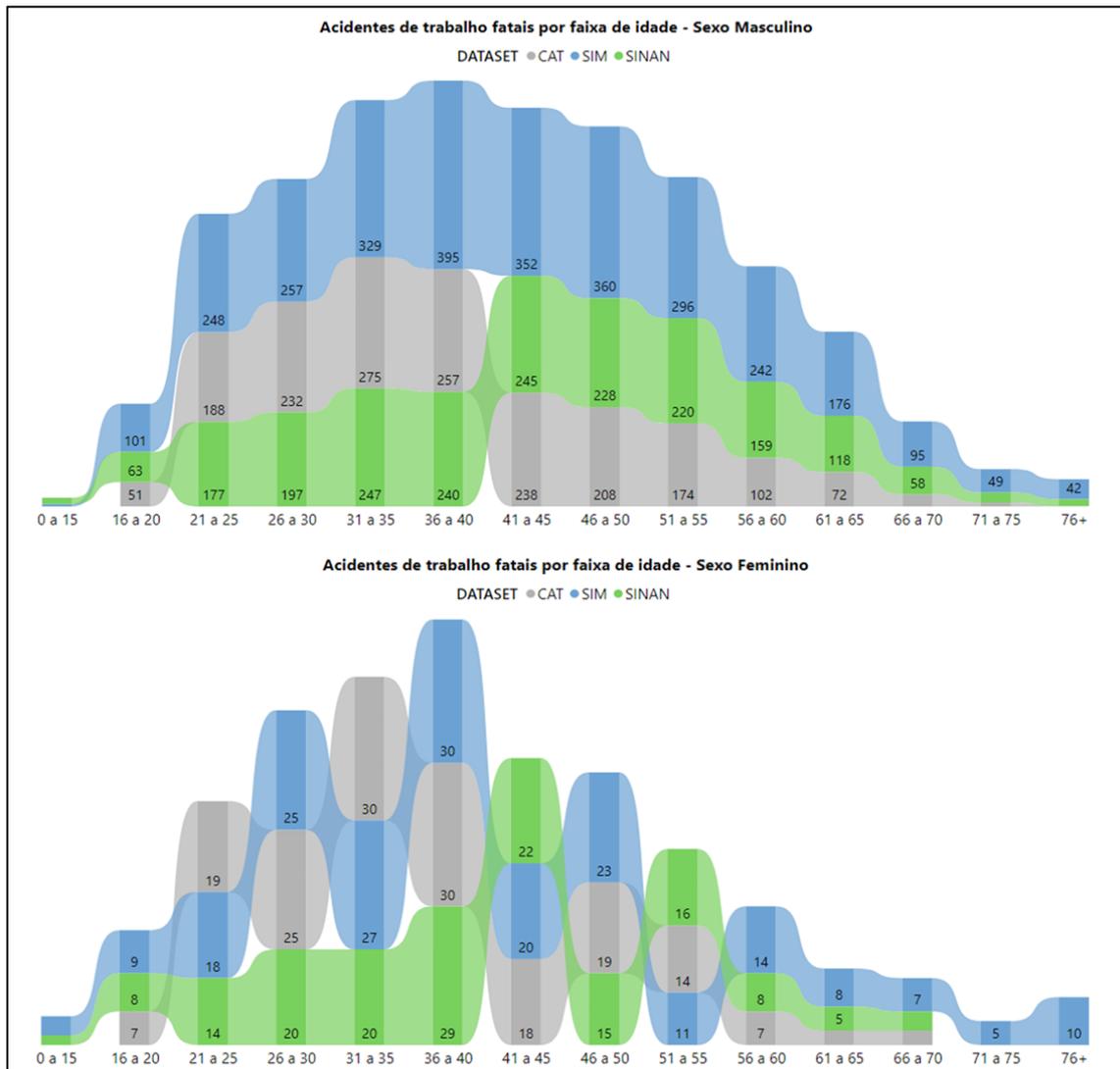
Tabela 3. Quantidade de acidente por sexo e fonte de dados

SEXO	CAT		SIM		SINAN	
	Quantidade	Percentual	Quantidade	Percentual	Quantidade	Percentual
Feminino	175	8,72%	211	6,68%	163	7,54%
Masculino	1.831	91,28%	2.947	93,32%	2.000	92,46%
Total	2.006	100,00%	3.158	100,00%	2.163	100,00%

Fonte: Autor, 2021

O afinamento do estudo da variável sexo na relação de acidente de trabalho por idade, conforme Figura 8, demonstra que a descrição dos casos gerais se aplica com alta similaridade com o sexo masculino. A população masculina com idade entre 21 e 55 anos estão mais suscetíveis a decorrência ao óbito quando ocorre um acidente de trabalho, destacando a faixa de idade entre 36 e 40 anos que representa o maior dos intervalos com letalidade de acidentes de trabalho. Quanto ao sexo feminino demonstra uma similaridade de comportamento com o sexo masculino com relação a acidentes fatais, com um pico de ocorrência entre os 31 e 35 anos de idade, e uma frequência relativa maior em comparação com a do sexo masculino.

Figura 8 - Distribuição dos registros por faixa etária, sexo e base de dados

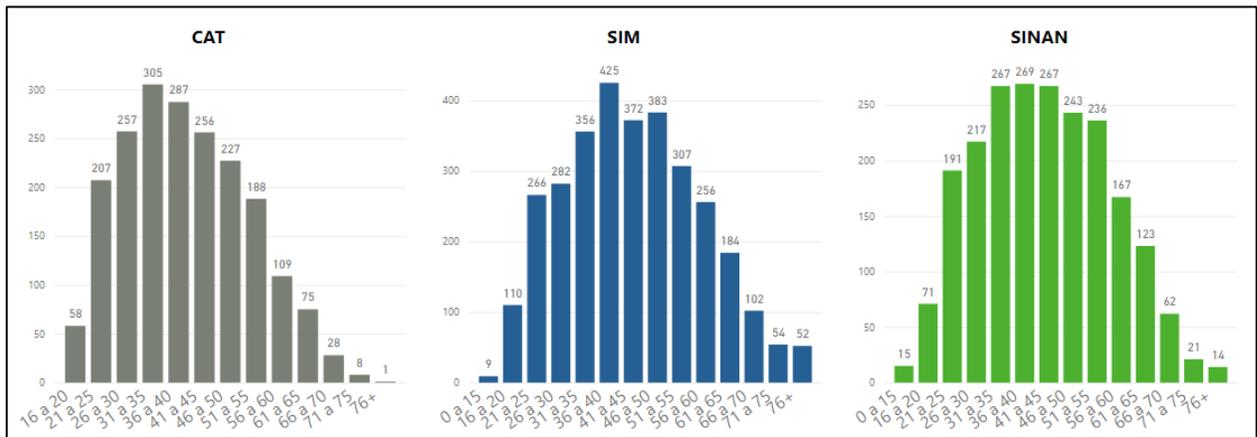


Fonte: Autor, 2021

Uma observação importante na Figura 8 é a influência da base em cada faixa de idade. Nota-se uma baixa de registro no banco de dados da CAT após os 60 anos de idade em comparação com as demais, com intensificação no sexo feminino. O comportamento apresentado na figura demonstra a necessidade de investigar os motivos que levam a subnotificação de dados nas principais bases voltada a registro de acidentes de trabalho.

A Figura 9 demonstra que conforme a população envelhece, o quantitativo de óbito tende a aumentar, começando a entrar em decaimento a partir da dos 45-50 anos. Este comportamento demonstra ser similar entre as bases, porém no SINAN o decaimento é um pouco mais tardio, após os 55 anos de idade.

Figura 9 - Quantidade de registro por faixa etária base de dados



Fonte: Autor, 2021

Voltando exclusivamente para os registros da CAT, pode-se notar que durante o ano de 2019 o setor econômico com maior quantidade de acidentes de trabalhos fatais concentrou-se na Indústria de Transformação, devido ao acontecimento de Brumadinho – MG, conforme comentado anteriormente. Porém destaca-se os setores de transportes, reparação de veículos automotores e construção civil.

Tabela 4 - Quantidade de acidentes fatais na CAT por classificação do CNAE

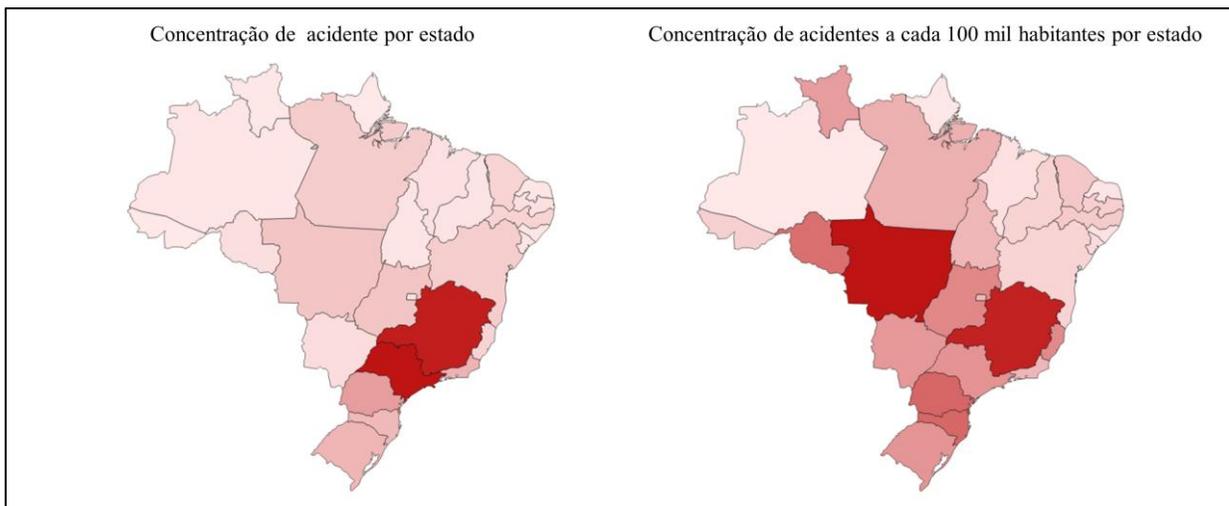
Denominação Seção	CAT	Percentual
Indústrias de transformação	387	18,61%
Transporte, armazenagem e correio	332	15,96%
Comércio; reparação de veículos automotores e motocicletas	306	14,71%
Construção	276	13,27%
Atividades administrativas e serviços complementares	160	7,69%
Indústrias extrativas	132	6,35%
Agricultura, pecuária, produção florestal, pesca e aquicultura	150	7,21%
Alojamento e alimentação	67	3,22%
Água, esgoto, atividades de gestão de resíduos e descontaminação	44	2,12%
Atividades profissionais, científicas e técnicas	44	2,12%
Administração pública, defesa e seguridade social	38	1,83%
Outras atividades de serviços	144	6,92%
Total	2.080	100,00%

Fonte: Autor, 2021

Abordando apenas alguns fatores e variáveis dos bancos de dados selecionados para esta pesquisa, o estudo de exploração dos dados permite mapear um panorama da situação atual. Conforme apresentada a quantidade de registros e a exploração de variáveis, demonstra-se a evidência da necessidade de estudar a subnotificação dentro dos bancos de dados que envolvem a saúde e segurança do trabalho. O mapeamento do cenário formado evidência que a subnotificação ocorre nos registros oficiais dos bancos de dados de acidentes fatais.

Geograficamente pode-se notar, conforme Figura 10, que espacialmente os acidentes de trabalho apresentam grande distinção quando visto por concentração, destacando o estado de São Paulo, o estado com a maior economia do país, seguido por Minas Gerais, que neste ano destacou devido ao desastre ambiental e humano de Brumadinho no interior de Minas Gerais. Porém quando relativiza os acidentes com a população de cada unidade federada, percebe-se que o estado do Mato Grosso ganha destaque com maior índice (2,3 acidentes fatais a cada 100 mil habitantes).

Figura 10 - Distribuição dos registros por unidade federada



Fonte: Autor, 2021

Os índices com a quantidade de acidente fatais em 2019 e a quantidade de acidente relativo a 100 mil habitantes pode-se verificar na Tabela 5. Mostrando o Sul como a primeira no índice relativo (1,31 acidentes por 100 mil habitantes) e em segunda posição em número absoluto (396 acidentes), seguido em segunda posição o Centro-Oeste no relativo (1,29 acidentes por 100 mil habitantes).

Tabela 5 - Quantidade de acidente fatal absoluto e relativo da CAT

Região / Estado	Acidente	Acidente por 100 mil hab.
Centro-Oeste	213	1,29
Distrito Federal	21	0,69
Goiás	83	1,17
Mato Grosso	81	2,30
Mato Grosso do Sul	28	1,00
Nordeste	239	0,42
Alagoas	9	1,29
Bahia	62	0,42
Ceará	49	0,53
Maranhão	21	0,30
Paraíba	19	0,47
Pernambuco	48	0,50
Piauí	14	0,43
Rio Grande do Norte	9	0,25
Sergipe	8	0,35
Norte	123	0,66
Acre	4	0,45
Amapá	2	0,23
Amazonas	9	0,21
Pará	66	0,76
Rondônia	25	1,39
Roraima	6	0,95
Tocantins	11	0,69
Sudeste	1109	1,25
Espírito Santo	46	1,13
Minas Gerais	459	2,16
Rio de Janeiro	122	0,70
São Paulo	482	1,04
Sul	396	1,31
Paraná	171	1,48
Rio Grande do Sul	118	1,03
Santa Catarina	107	1,48
Total Geral	2080	0,98

Fonte: Autor, 2021

Importante ressaltar que a localidade na CAT é o da empresa que o empregador trabalha e não necessariamente o local do acidente. Assim acidentes como por exemplo do meio de logística pode mudar o cenário apresentado. Tomou-se esta variável como escolha para fazer análise, devido a variável de UF do acidente estar incompleta para alguns estados brasileiros.

Abordando apenas alguns fatores e variáveis dos bancos de dados selecionados para esta pesquisa, o estudo de exploração dos dados permite mapear um panorama da situação atual. Conforme apresentada a quantidade de registros e a exploração de variáveis, demonstra-se a evidência da necessidade de estudar a subnotificação dentro dos bancos de dados que envolvem a saúde e segurança do trabalho. O mapeamento do cenário formado evidência que a subnotificação ocorre nos registros oficiais dos bancos de dados de acidentes fatais.

O próximo capítulo consiste na aplicação do método de captura e recaptura, por meio do estimador de Lincoln-Petersen para determinar a subnotificação de casos entre as fontes de dados.

5

ANÁLISE DE RESULTADOS

A exploração das variáveis por meio de *data visualization* dos *datasets* da CAT, SINAN E SIM contribuiu para um melhor entendimento da distribuição das informações por meio de análises espaciais e temporais. Pôde-se perceber que comportamentos isolados e a similaridade entre os *datasets* conduz a uma série de evidências de subnotificação de acidentes fatais dentro das plataformas que registram estas informações. A quantidade de registro e a relação das variáveis conforme abordados no capítulo anterior corroboram estas evidências.

Do ponto de vista Saúde e Segurança do Trabalho, a subnotificação dentro da previdência Social (não abertura dos registros da CAT), são os que apresentam maior impacto nos direitos dos envolvidos, neste caso os familiares. Porém estende-se também para controle de políticas dos direitos e saúde e segurança do trabalhador. Os direitos envolvem a formalidade dos empregos ao longo do território nacional, visto que a formalização de carteira assinada não é homogênea no território nacional. Com relação a saúde e segurança do trabalhador, o impacto está diretamente ligado a integridade da vida da pessoa, uma vez que o no caso extremo do acidente, é irreversível a consequência da fatalidade.

Portanto mensurar subnotificação entre as plataformas de registros de acidentes de trabalho torna-se importante para desenvolver diretrizes e políticas na prevenção de acidentes de trabalhos e direcionar para comportamento que se tornam marcantes na fatalidade nesta classe de acidentes. Embora a abertura de comunicado de acidente de trabalho – CAT – esteja regulamentado na lei

existe outras plataformas que realizam o registro ou apontamento de acidentes relacionados ao trabalho que podem auxiliar na determinação da subnotificação. A plataforma do DATASUS, neste caso os *datasets* do SINAN e o SIM, registram a relação de acidentes graves relacionados ao trabalho. Como os *datasets* não apresentam informações suficiente que possa identificar os indivíduos em comum entre os bancos de dados, necessita-se, portanto, de meios que auxiliem na identificação dos casos para determinar a subnotificação de dados por meio de ferramentas estatísticas.

5.1 DETERMINAÇÃO DAS VARIÁVEIS

A identificação das variáveis em comum entre os *datasets* – CAT, SINAN e SIM – constitui o primeiro passo para uma posterior comparação e identificação de similaridade de casos. A Tabela 6, mostra que diante das inúmeras variáveis presente em cada *dataset*, pode-se reduzir a uma relação de 7 variáveis.

Tabela 6 - Classificação das variáveis dos *Datasets*

Variável	Tipo	Datasets		
		CAT	SINAN	SIM
Sexo	Catégorica	✓	✓	✓
Data de Nascimento ou Idade	Quantitativa*	✓	✓	✓
Data do Acidente	Quantitativa*	✓	✓	✗
Data do Óbito	Quantitativa*	✗	✓	✓
Abertura de CAT	Catégorica	✓	✗	✗
Local do Acidente	Catégorica	✓	✓	✓
CID-10	Catégorica	✓	✓	✓

* - quando considerado intervalo de tempo, como por exemplo: idade e tempo do acidente.

Fonte: Autor, 2021

Verificando se as variáveis apresentavam consistência entre os 3 *datasets*, pode-se notar que no caso das geográficas, como por exemplo UF e Município, não tinha uma relação direta entre a CAT e o SINAN, devido à uma alta taxa de ausência de dados para regiões do sul, sudeste e centro-oeste brasileiro na base de dados da CAT. Assim descartou a variável regional dos estudos que envolvesse a CAT com os demais *datasets*. O CID-10 embora seja comum nos três *datasets*, apresentou uma relação de baixa confiabilidade, possivelmente a parte humano no preenchimento da informação e não dependência das informações entre as fontes de dados.

A próxima etapa é primordial para a determinação da subnotificação de acidentes fatais com base nos *datasets* utilizados, consistindo na descrição do procedimento utilizado para determinação dos casos em comum nos diferentes bancos de dados.

5.2 TRATAMENTO DOS DADOS

Com as bases padronizadas e a determinação das variáveis que seriam utilizadas para gerar a quantidade de casos em comum nos arranjos dos *datasets*, realizou-se análises isolada dentro de cada um para verificar o comportamento e o padrão de cada variável.

Diante da existência de um número expressivo de variáveis isoladas em cada *dataset*, a quantidade de variáveis em comum entre as fontes de dados é relativamente baixa, conforme apresentado no Tópico 5.1. Um ponto importante a ser considerado é o nível de confiabilidade entre as fontes de dados, porém apenas com a limitação das informações dentro de cada uma, não é possível determinar esta qualidade de informação entre os data sets.

A relação de várias em comum entre os *datasets* da CAT e SINAN limitou-se apenas a:

1. Sexo: Variável categórica
2. Data de Nascimento ou idade: Variável quantitativa quando considerado idade;
3. Data do Acidente: Variável quantitativa quando considerado idade

Após o tratamento dos dados, realizou a integração dos 3 *datasets*, em outros dois diferentes *datasets*, sendo: [1] CAT-SINAN e [2] CAT-SIM. Após a junção dos respectivos arquivos realizou-se a concatenação das variáveis para cada atributo (caso), gerando uma espécie de código, que corresponderia a cada indivíduo analisado. Neste momento foi aplicada uma das premissas, de quando houver mais de um caso em comum dentro da mesma fonte de dados, eles seriam interpretados como distintos.

Com cada *dataset* pronto, verificou quantos casos de origem da fonte de dados 1 estava presente na 2. Obtendo assim de forma simplificada o *input* necessário para determinação da subnotificação de acidente de trabalhos fatais.

5.3 PREMISSAS

Diante dos critérios do modelo de Captura e Recaptura, considerou-se algumas premissas importantes para a continuidade no estudo, sendo:

1. Que os casos que apresentam relação entre os *datasets* refere-se realmente ao mesmo indivíduo trabalhador;
2. Quando dentro de um mesmo *dataset* houver duplicatas de casos, que caracteriza um indivíduo, ele será considerado como casos distintos. Importante destacar que pode ocorrer preenchimento duplicados principalmente no caso da CAT, devido não ter um responsável único na abertura do comunicado de acidente de trabalho;
3. Devido limitação temporal, será considerado que dentro do ano de estudo (2019) todos os casos foram devidamente abertos. Tem-se ciência que principalmente no caso do SINAN e do SIM, pode haver atraso na notificação dos casos, ocorrendo uma subnotificação temporal dentro dos próprios bancos.

5.4 REQUISITOS CAPTURA E RECAPTURA

Este tem como finalidade garantir a aplicabilidade da técnica de captura e recaptura para determina a subnotificação de acidentes de trabalhos fatais entre as fontes de dados da Previdência Social e do Sistema Único de Saúde. Assim tomou os devidos cuidados para garantir que:

- A população seja fechada: entende-se como população fechada aquela na qual não são observados nascimentos, mortes ou migrações durante o período de estudo. Embora este estudo atua na base de óbitos não existe nenhuma ação na qual possa interferir na captura dos indivíduos. O fato de atuar sobre bases já consolidadas auxilia na corroboração desse requisito;
- Tenha uma identificação perfeita dos indivíduos: esse tópico é responsável por um dos maiores desafios do presente trabalho que consiste na identificação perfeita dos indivíduos. Como nenhuma base apresenta uma relação direta com as demais, é necessário a utilização de métodos que permitisse garantir que a identificação entre os *datasets* fosse estabelecida da melhor forma possível;
- Que as amostras sejam independentes: a independência das amostras é garantida por meio dos registros nos diferentes *datasets*, visto que nenhum tem um vínculo de dependência em relação ao outro. Assim pode ser garantir que todos os registros nas fontes de dados estão independentes um do outro;

- Que em cada amostra a probabilidade de seleção seja constante: neste tópico a garantia de probabilidade de seleção ser constante está não utilização completa dos bancos de dados, ou seja, trabalhar com a amostra completa com o tamanho populacional e não amostral.

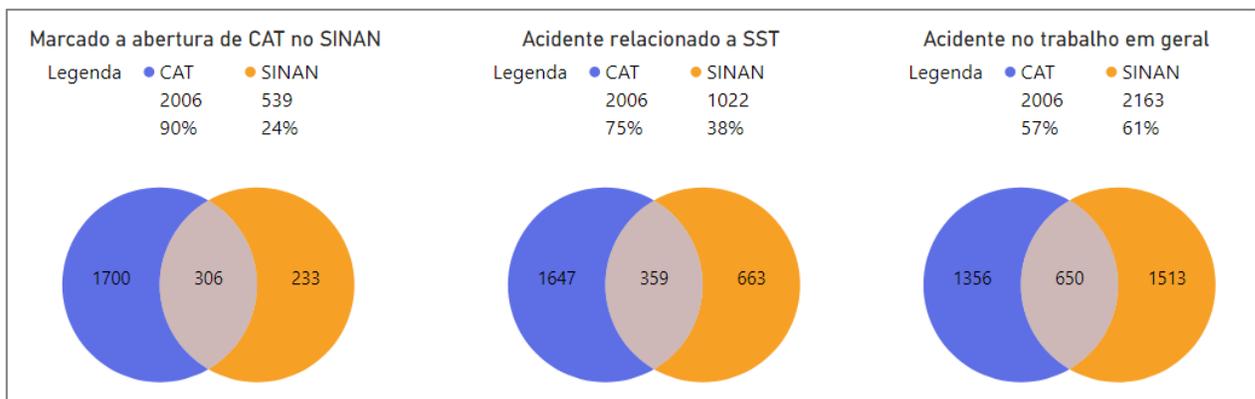
5.5 CAT E SINAN

Os *datasets* CAT e SINAN apresentam um maior nível de contribuição voltado para acidentes de trabalho visto que ambos têm foco central para esta categoria de acidentes. Embora as variáveis de cada um dos *datasets* apresenta certas limitações de relação um com o outro, ainda assim é possível realizar um estudo com base nestas informações.

Uma variável importante dentro da fonte de dados do SINAN está na indicação se para cada registro houve a abertura de um Comunicado de Acidente de Trabalho – CAT – para a ocorrência. Esta informação é muito importante para determinar o nível de subnotificação dentro da base e podendo obter mais resultados no estudo. Diante desta variável obteve-se três casos de análise:

- Caso 1: Com a marcação de abertura de um Comunicado de Acidente de Trabalho (CAT) no SINAN;
- Caso 2: Os casos com marcação de CAT adicionado os casos de acidente de trabalho, porém sem abertura de CAT;
- Caso 3: além dos casos anteriores, estão inclusos todos os acidentes que têm alguma relação com trabalho, porém não necessariamente com prestação de serviço.

Figura 11 - Diagrama de Venn entre CAT e SINAN



Fonte: Autor, 2021

O Diagrama de Venn denotado na Figura 11 apresenta graficamente a distribuição dos conjuntos de dados, conforme os casos apresentados no parágrafo anterior. A figura permite observar como está estabelecida a relação entre os *datasets*.

A Tabela 7 apresenta os parâmetros utilizados na aplicação do cálculo de estimação da subnotificação de acidentes fatais no ambiente de trabalho por meio do estimador de Lincoln-Petersen. Aplicou-se também o estimador de Chapman, porém o resultado apresentou muita proximidade com o Lincoln-Petersen, no qual foi realizada a análise.

Tabela 7 - Subnotificação de acidentes fatais CAT e SINAN

Casos	CAT (N_{A1})	SINAN (N_{B1})	CAT ∩ SINAN (n₁₁)	E-LP (N)	E-CH (N)	Subnotificado (n₂₂) – E-LP	Subnotificação E-LP
Caso1	2.006	539	306	3.533	3.529	682	19%
Caso2	2.006	1.022	359	5.711	5.702	2.324	41%
Caso3	2.006	2.163	650	6.675	6.671	1.856	28%

Legenda: E-LP: Estimador Lincoln-Petersen; E-CH: Estimador de Chapman.

Fonte: Autor, 2021

O primeiro registro refere-se a acidentes de trabalho fatais no qual o SINAN indicou a abertura de uma CAT para a ocorrência. Contudo a quantidade de indivíduos recapturados ($CAT \cap SINAN = 306$) indica que nem toda a população do SINAN está presente na CAT, levando a uma taxa de subnotificação de 19% na população total e de 43,2% na base da Previdência Social – CAT. O segundo caso da Tabela 7 evidencia a subnotificação uma vez que marca as ocorrências no qual não houve a abertura de um Comunicado de Acidente de Trabalho, levando a um expressivo número de 41% de subnotificação geral e de 64,9% na base de dados da CAT. O último caso embora apresenta uma taxa de subnotificação de 28% está incluso também acidentes que envolvem acidentes domésticos de trabalho, como por exemplo: Um pedreiro que sofreu acidente trabalho na atividade em casa e outros casos correlatos, sem que possa haver uma relação empregador-empregado, assim quando considerado a abertura de CAT, a subnotificação dos acidentes aproxima de 70% na base da previdência social.

5.6 CAT E SIM

Diferentemente do SINAN, a base de dados do Sim não apresenta muitas informações referente ao tem ade acidente de trabalho ocupacional, devido a natureza das ocorrências. O SIM

apresenta os registros de óbitos que ocorrem dentro da estrutura do SUS, assim os dados estão muito mais relacionados a saúde do indivíduo do que com a forma que levou a fatalidade.

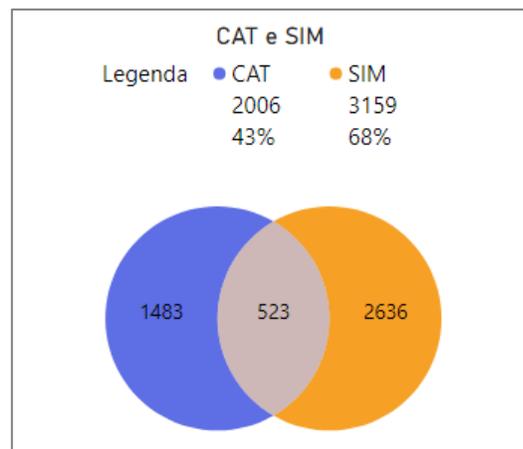
A ausência de um relacionamento maior com a temática de acidente de trabalho, pode aumentar a imprecisão do estudo, visto que

Da mesma forma que realizou na análise da CAT com o SINAN, na análise do SIM relação de várias em comum entre os *datasets* da CAT e SINAN limitou-se apenas a:

1. Sexo: Variável categórica
2. Data de Nascimento ou idade: Variável quantitativa quando considerado idade;
3. Data do Acidente: Variável quantitativa quando considerado idade;

De mesma forma do anterior o Diagrama de Venn denotado na Figura 12 apresenta graficamente a distribuição dos conjuntos de dados entre os *datasets* da CAT e do SIM.

Figura 12 - Diagrama de Venn entre CAT e SIM



Fonte: Autor, 2021

Em posse de todos os *inputs* necessário, fez-se uso do estimador de Lincoln-Petersen (Eq. 1) para determinar a subnotificação conforme Tabela 8. De forma análoga aplicou-se também o estimador de Chapman, porém o resultado apresentou muita proximidade com o Lincoln-Petersen, no qual foi realizada a análise.

Tabela 8 - Subnotificação de acidentes fatais CAT e SIM

CAT (N_{A1})	SINAN (N_{B1})	CAT \cap SINAN (n_{11})	E-LP (N)	E-CH (N)	Subnotificado (n_{22})	Subnotificação
2.006	3.159	523	12.116	12.102	6.428	53%

Fonte: Autor, 2021

A determinação da subnotificação do banco de dados da CAT em decorrência com o SIM obteve-se um grau de similaridade no caso 3 do Sinan, no qual estavam inclusos os casos gerais, seja com abertura ou não da CAT, levando a uma subnotificação de 53% dos acidentes de trabalho fatais em ambiente de trabalho. Dos 12.116 casos estimados de fatalidades acontecidas em ambiente organizacional de trabalho, apenas 2.006 abriram um comunicado de acidente de trabalho (CAT), representando um elevado percentual de 83%.

5.7 DISCUSSÃO DOS RESULTADOS

Os resultados apresentaram o grande desafio de trabalhar com as fontes de dados públicos brasileira. Da seleção dos registros a aplicação de técnicas que permitem estudar o comportamento dos registros são encontrados diversos obstáculos para o avanço da pesquisa.

Diante da limitação das variáveis com similaridade entre as fontes de dados, necessitou um tratamento dos registros para capturar o máximo de dados possível, realizando limpeza nas variáveis e retirando atributos (casos) que apresentam falta (*missing*) de informações importantes para a condução dos estudos.

As variáveis de datas no geral têm classificações desconhecidas, tomando características apenas quando é possível estabelecer uma relação de tempo para o evento, como idade, tempo do acidente ou tempo do registro de acidente de trabalho. Como são sensíveis ao registro, estes campos têm margem de erros, que foram encontrados durante o estudo, no preenchimento dos formulários. Em especial a variável que determinaria a idade, verificou-se que devido a possíveis falhas de informações no cadastramento da data do acidente, que determinaria a idade que a pessoa tinha até sofrer o acidente que levou ao óbito, existia também uma possibilidade dessas datas terem sido cadastrado erradas. Um possível motivo para este comportamento dá-se que nem sempre a regulamentação de fazer o registro com até 24 horas é seguida.

Durante o processo de identificação das variáveis em comum presente nos *datasets* deparou-se com alguns desafios que pode influenciar nos resultados dos estudos:

1. Confiabilidade dos dados: um ponto primordial de qualquer estudo que envolva dados é a qualidade e confiabilidade de dados. A importância desta característica está relacionada a nível de decisão estratégica a ser tomada com o resultado do

estudo. Como os *datasets* deste estudo estão sujeitos a subjetividade de preenchimento por parte humana, classificações e erros podem comprometer a qualidade da informação inseridas. Erros como por exemplo: data de nascimento errada e fora dos padrões, que fogem a realidade e o contexto de acidente de trabalho, como nascimento em séculos seguintes ou então com idades fora da longevidade humana. No contexto da subjetividade está na classificação de um acidente dentro do CID-10, que embora seja padronizado, pode apresentar divergência no preenchimento por diferentes pessoas, dependendo dos cenários do acidente.

2. Consistência das informações: entende-se como consistência das informações se as informações preenchidas em um determinado *dataset* serão equivalentes ao outro. Caso não haja uma verdadeira equivalência o erro do estudo tende-se a aumentar;
3. Integração dos dados: Como os dados são provenientes de diferentes plataformas e de origem governamental, a disponibilização e frequência das informações estão na dependência de cada setor fornecer em suas respectivas plataformas.

Com as variáveis determinadas e padronizadas, estabeleceu-se os índices de subnotificações para cada relação de fonte de dados, e algumas considerações importantes foram extraídas juntamente com o resultado da aplicação do estimador. O foco da pesquisa está centrado na CAT, devido ser o meio de notificação que garante os direitos do trabalhador, sendo assim a determinação da subnotificação aconteceu por meio da relação da CAT-SINAN e CAT-SIM. A relação SIM x SINAN não está alinhada com o objetivo do trabalho, sendo desconsiderada para análise.

O presente estudo apresenta alguns aspectos que interferir no nível da acuracidade dos resultados dos estimadores.

1. O primeiro aspecto é devido a não identificação perfeita dos casos entre os diferentes *datasets*, uma vez que não tem presente o número de CPF ou variáveis que indicam uma direta relação do acidente com o trabalhador. Para garantir que os casos fossem devidamente assimilados, fez-se o uso do maior número possível de variáveis para criar uma identificação de cada caso entre os *datasets*, estando sujeito a confiabilidade da técnica utilizada neste trabalho.

2. O estudo também pode estar ameaçado quanto a confiabilidade das informações dentro de cada fonte de dados, neste caso a qualidade dos registros e se apresentam a integridade das informações que conforme foi abordado durante o estudo apresentou registros e evidências de ausência de informações e erros no registro das ocorrências.
3. A subnotificação temporal dentro da própria base é um aspecto que pode comprometer o nível de similaridade entre a realidade e a estimativa do estudo, uma vez que bases como o SIM e o SINAN apresentam indício forte de registro de óbitos em anos posteriores.

A Seção seguinte pode ser conferida algumas considerações importantes da determinação da subnotificação.

5.7.1 Subnotificação CAT x SINAN

A determinação da subnotificação por meio da relação de bases CAT e SINAN apresentou índices importantes, visto que ambas as fontes de dados têm foco voltado para a Saúde e Segurança do Trabalhador.

O Caso 1, que consiste na subnotificação de acidentes de trabalho fatais que apresentam marcação de abertura de CAT dentro da base de dados do SINAN, apresentou algumas considerações importantes. Partindo da premissa que para este caso: todos os indivíduos estariam presente na fonte de dados da CAT, alguns possíveis motivos que levariam a esta taxa de subnotificação, que neste caso é considerada alta seria:

- Erro no preenchimento das informações nas fichas, levando a divergência de informações entre as fontes;
- O SINAN considerou como indicação de abertura de CAT, o comunicado de acidente apenas, visto que a abertura de acidente e de óbito são diferentes, e nem sempre o óbito acontece de forma imediata.
- Temporalidade da distribuição das informações, visto que enquanto os relatórios com as informações da CAT são disponibilizados trimestralmente os do SINAN leva um período de um ano para disponibilizar.

- Disponibilização completa das ocorrências registradas no SINAN, visto que é uma plataforma mais recente de registro de informações relacionadas ao trabalho e nem todas as informações podem estar sendo passadas para esta fonte de registros.

No Caso 2, que inclui o Caso 1 adicionado os registros de acidentes de trabalho que não apresentaram abertura de CAT para a fatalidade. Neste caso, além dos motivos elencados anteriormente, pode-se considerar aqui a informalização da mão de obra. A seguridade do acidentado está ligada diretamente a formalização do emprego, tornando para aqueles que não apresentam carteira de trabalho assinada, um processo mais lento e árduo para requerimento dos direitos dos envolvidos.

O Caso 3, que apresenta o Caso 1, Caso 2 e os demais registros ligados a alguma forma de trabalho apresentou uma subnotificação inferior ao Caso 2, 53% do Caso 2 e 47% do Caso 2. Esperava-se uma inversão da subnotificação destes casos, ou seja, que o Caso 2 a subnotificação fosse menor que no Caso 3. Uma possível justificativa para este comportamento pode estar na identificação dos casos, ou seja, que a quantidade de registros em comum entre as fontes de dados não são necessariamente os mesmos registros.

Mesmo diante das limitações e a confiabilidade das considerações colocadas para condução deste trabalho, a subnotificação de acidentes de trabalho fatais torna-se alta, fazendo surgir problemas de seguridade, direito e saúde do trabalhador. É de se esperar que a fonte de dados do SINAN tivesse maior consistência com a CAT, visto que existe uma relação de informação entre ambas, porém o estudo mostra que há uma grande lacuna de informações entre ambas. Considerando que a legislação é mais rígida quando ocorre a fatalidade do acidente, é de se esperar que para acidentes de trabalho não fatais a subnotificação seja no mínimo igual ou maior.

5.7.2 Subnotificação CAT x SIM

De forma similar ao Caso 3 do SINAN, a determinação da subnotificação por meio das fontes de dados da CAT e do SIM, pode não apresentar necessariamente uma relação de trabalho do tipo empregado-empregador, assim a similaridade dos casos pode-se ocorrer devido a forma de notificação dos casos dentro do SUS, uma vez que ocorre apenas a marcação que o acidente de trabalho fatal tem uma certa relação com acidentes de trabalho.

Portanto coloca-se em questionamento de que a fatalidade ocorrida está realmente relacionada a um acidente de trabalho organizacional, ou seja, o óbito pode ser decorrente de um acidente domiciliar por exemplo.

Uma observação muito importante também é que por meio do SUS pode-se ter registro de acidentes de trabalhos não formalizados na carteira de trabalho, nestes casos espera-se realmente que a subnotificação seja maior uma vez que dependendo da situação e das características do trabalhador pode se tornar difícil a abertura para requerimento dos direitos garantidos por lei, porém submetido a uma realidade ainda forte em território nacional.

6

CONCLUSÃO

A exploração das fontes de dados da CAT, SINAN E SIM contribuiu para entender melhor como é a relação dos registros de acidentes de trabalhos entre estas fontes, levando a expressivos números de acidentes fatais subnotificados, porém abrindo campo além da pesquisa científica, para tomadores de decisão envolvidos na temática. A seguir é apresentado um fechamento da pesquisa e as contribuições do estudo.

6.1 CONSIDERAÇÕES FINAIS

Por meio dos estimadores de Lincoln Petersen e Chapman, determinou-se as taxas de subnotificação de acidentes fatais no Brasil entre as fontes de dados públicas da CAT, SIM e SINAN. Os resultados respondem à questão de pesquisa, sendo interpretados com criticidade dentro do cenário de acidentes de trabalhos fatias em território nacional.

Mediante ao uso de *data visualization*, possibilito o estudo exploratório, espacial e temporal das variáveis nos bancos de dados, evidenciado possíveis subnotificações de acidentes de trabalhos. A condução da pesquisa resultou na determinação nos números e taxas de subnotificação de acidentes de trabalhos diante o relacionamento das fontes de dados da CAT, SINAN e SIM.

Antes de obter os resultados do trabalho, grandes desafios foram deparados devido as dificuldades de trabalhar com fonte de dados públicas, já que apresentam inconsistências de

informações, falta de confiabilidade e qualidade nos registros e uma baixa relação de variáveis em comum entre as diferentes fontes, mesmo as que apresentam uma relação mais direta a acidentes de trabalho, como o caso da CAT e SINAN.

Resultados mostraram que a relação entre as fontes de dados apresenta alto índice de subnotificações. A relação de registros entre a CAT e o SINAN apresentou uma taxa de subnotificação de 43%, mesmo diante da confirmação da abertura de comunicação de acidente de trabalho dentro do SINAN, que pode ser justificadas pela inconformidade dos dados, disponibilidade dos registros e tipo de comunicado de acidente de trabalho aberto na previdência social. Outro número muito importante está nos 65% de subnotificação e acidentes de trabalhos que apresentam ou não abertura de CAT na previdência social, que pode estar associado a trabalho informal no mercado agravando para outra importante definição do ambiente econômico.

A subnotificação de acidentes de trabalhos fatais quando comparado o banco de dados da CAT com o SIM apresenta uma taxa de 53%. Para registros gerais relacionado a trabalho, não necessariamente com algum vínculo empregatício, a subnotificação da CAT em relação ao SINAN apresenta uma relação parecida com a CAT em relação ao SIM, estimando uma subnotificação de 70% e 83% respectivamente. Trabalho informal, acidentes sem relação de prestação de serviço, erro de preenchimento, entre outros, são possíveis motivos que podem elevar estas taxas e apresentar um cenário grave de acidente de trabalho fatal, levando o estudo para possíveis problemas sociais.

Mesmo diante de todas as limitações, o presente trabalho tem grande contribuição para o meio acadêmico, por meio da exploração de dados públicos com técnicas estatísticas, porém também para ou auxílio à tomadores de decisões relacionado à saúde e segurança do trabalho. Embora o efeito tempo tenha grande importância na relevância do estudo, a realização de análises em tempo mais atual a pesquisa pode gerar grandes diretrizes para as esferas sociais e governamentais envolvidas. Assim será possível desenvolver políticas de saúde e segurança do trabalho além de dar direcionamentos para investigações e fiscalizações nos setores mais incidentes ou regiões com maiores índices de fatalidades no ambiente de trabalho.

6.2 LIMITAÇÕES DO ESTUDO

O presente estudo foi de grande desafio, iniciando pelo fato de trabalhar com fonte de dados governamentais, que nem sempre são construídas voltadas para exploração e criação de conteúdo a partir dos dados.

Um dos maiores desafios de trabalhar com dados é justamente a capacidade de obter insights importante e relevantes para construção de estudos. Embora o presente trabalho apresente um foco na subnotificação, considerou-se ao longo de todo o trabalho se este estudo poderia ter desdobramento por meio de outras variáveis como região, agente causador, setor da econômica. Porém dadas as limitações das relações entre as fontes de dados, não foi possível conduzir o estudo a este nível de detalhamento.

Quanto a abordagem da temporalidade dos eventos de acidentes de trabalho no desenvolvimento do estudo, torna-se obstáculo devido a disponibilização por cada fonte de dados seja do DATASUS ou DATAPREV serem de constância e períodos diferentes. Com relação a este tópico é possível estabelecer uma relação direta quanto a relevância do assunto abordado, uma vez que quanto mais antigo é a disponibilização dos dados, menos relevantes as informações se tornam para aplicação de decisões imediatas. Considerando a globalização e as rápidas mudanças que a economia passa constantemente, é fundamental que para certas decisões obtenha as informações o mais recente possível. Uma disponibilização mais ágil dos dados, conduzirá a realização de estudos e análises, com objetivo de fornecer direcionamento relevantes para os tomadores de decisões, para desenvolvimento de políticas de saúde e segurança pública além de direcionar nas ligações nas investigações e fiscalizações no setor os setores envolvidos.

Com relação a confiabilidade das informações, também foi um ponto importante para esta pesquisa. Diante da baixa qualidade de alguns registros e a confirmação da veracidade das informações dentro de cada fonte de registro de acidente, realizou alguns procedimentos para remoção destes dados, porém de forma que impactasse o mínimo possível na pesquisa. A relação das informações entre fontes de dados também foi considerada sobre este viés, visto que o preenchimento das informações nos registros de acidente pode estar suscetível a subjetividade e erro da pessoa responsável.

Uma limitação muito importante para este estudo e o maior desafio se caracterizou por meio da identificação da relação dos casos entre as diferentes fontes de dados. A ausência de variáveis que poderiam auxiliar na determinação de casos equivalentes, conduziu a uma busca por métodos

que pudesse auxiliar nesta etapa e fornecer maior confiabilidade e segurança no andamento da pesquisa.

6.3 TRABALHOS FUTUROS

O presente trabalho deixou alguns questionamentos a serem respondidos em pesquisas futuras, diante das limitações das fontes de dados utilizados. Encontrar novas fontes de dados de acidentes de trabalho ou até mesmo obter informações mais completas das fontes utilizadas pode fornecer grandes direcionamentos a novas pesquisas e responder perguntas que na presente pesquisa ficou na espera de respostas como:

- Q1: Quais são os fatores que influenciam na ocorrência de acidentes de trabalho? E como estes fatores podem determinar a subnotificação?
- Q2: Quais são as tendências identificadas na ocorrência de acidentes de trabalho? Como o nível de acidentes fatais subnotificados pode influenciar em medidas macro e micro da prevenção de acidentes de trabalho?

REFERÊNCIAS

ABEGAZ, Teferi et al. Road traffic deaths and injuries are under-reported in Ethiopia: a capture-recapture method. **PLoS One**, v. 9, n. 7, 2014. DOI: 10.1371/journal.pone.0103001. Disponível em: <https://doi.org/10.1371/journal.pone.0103001>.

AHMED, Ashar; SADULLAH, Ahmad Farhan Mohd; YAHYA, Ahmad Shukri. Errors in accident data, its types, causes and methods of rectification-analysis of the literature. **Accident Analysis & Prevention**, v. 130, p. 3-21, 2019. DOI: 10.1016/j.aap.2017.07.018. Disponível em: <https://doi.org/10.1016/j.aap.2017.07.018>.

AIT-MLOUK, Addi; GHARNATI, Fatima; AGOUTI, Tarik. An improved approach for association rule mining using a multi-criteria decision support system: a case study in road safety. **European transport research review**, v. 9, n. 3, p. 40, 2017. DOI: 10.1007/s12544-017-0257-5. Disponível em: <https://doi.org/10.1007/s12544-017-0257-5>.

ALBERT, Alex; PANDIT, Bhavana; PATIL, Yashwardhan. Focus on the fatal-four: implications for construction hazard recognition. **Safety science**, v. 128, p. 104774, 2020. DOI: 10.1016/j.ssci.2020.104774. Disponível em: <https://doi.org/10.1016/j.ssci.2020.104774>.

AMSTRUP, Steven C.; MCDONALD, Trent L.; MANLY, Bryan FJ (Ed.). Handbook of capture-recapture analysis. Princeton University Press, 2010.

AZAROFF, Lenore S.; LEVENSTEIN, Charles; WEGMAN, David H. Occupational injury and illness surveillance: conceptual filters explain underreporting. **American journal of public health**, v. 92, n. 9, p. 1421-1429, 2002. DOI: 10.2105/ajph.92.9.1421. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1447253/>.

BARBOSA, A. et al. Subnotificação de acidente ocupacional com materiais biológicos entre profissionais de Enfermagem em um hospital público. **Rev Bras Med Trab**, v. 15, n. 1, p. 12-7, 2017. DOI: 10.5327/Z1679443520177034. Disponível em: <https://cdn.publisher.gn1.link/rbmt.org.br/pdf/v15n1a03.pdf>.

BATISTA, Adriana Galdino; SANTANA, Vilma Sousa; FERRITE, Silvia. Registro de dados sobre acidentes de trabalho fatais em sistemas de informação no Brasil. **Ciência & Saúde Coletiva**, v. 24, p. 693-704, 2019. DOI: 10.1590/1413-81232018243.35132016. Disponível em: <https://doi.org/10.1590/1413-81232018243.35132016>.

BEHERA, Ramesh Kumar; HASSAN, Md Izhar. Regulatory interventions and industrial accidents: A case from India for 'Vision Zero' goals. **Safety science**, v. 113, p. 415-424, 2019. DOI: 10.1016/j.ssci.2018.12.013. <https://doi.org/10.1016/j.ssci.2018.12.013>.

BENAVIDES, Fernando G. et al. Comparison of fatal occupational injury surveillance systems between the European Union and the United States. **American journal of industrial medicine**, v. 44, n. 4, p. 385-391, 2003. DOI: 10.1002/ajim.10290. Disponível em: <https://doi.org/10.1002/ajim.10290>.

BEVILACQUA, Maurizio; CIARAPICA, F. E.; GIACCHETTA, G. Industrial and occupational ergonomics in the petrochemical process industry: A regression trees approach. **Accident Analysis & Prevention**, v. 40, n. 4, p. 1468-1479, 2008. DOI: 10.1016/j.aap.2008.03.012. Disponível em: <https://doi.org/10.1016/j.aap.2008.03.012>.

BORDONI, Polyanna Helena Coelho et al. Utilização do método de captura-recaptura de casos para a melhoria do registro dos acidentes de trabalho fatais em Belo Horizonte, Minas Gerais, 2011. **Epidemiologia e Serviços de Saúde**, v. 25, p. 85-94, 2016. DOI: 10.5123/S1679-49742016000100009. Disponível em: <https://doi.org/10.5123/S1679-49742016000100009>.

BOUCAUT, Rose A.; KNOBBEN, Linda M. Occupational health and safety incidents at a university school of nursing: A retrospective, descriptive appraisal. **Nurse education in practice**, p. 102776, 2020. DOI: 10.1016/j.nepr.2020.102776. Disponível em: <https://doi.org/10.1016/j.nepr.2020.102776>.

BRASIL. Decreto-Lei 5.452 de 1º de maio de 1943. **Consolidação das Leis do trabalho**. Disponível em: http://www.planalto.gov.br/ccivil_03/decreto-lei/del5452.htm. Acesso em: 25 set. 2021.

BRASIL. Lei nº 12.527, de 18 de novembro de 2011. **Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal; altera a Lei nº 8.112, de 11 de dezembro de 1990; revoga a Lei nº 11.111, de 5 de maio de 2005, e dispositivos da Lei nº 8.159, de 8 de janeiro de 1991; e dá outras providências**. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/112527.htm. Acesso em: 25 set. 2021.

BRASIL. Lei Nº 8.213, de 24 de julho de 1991. **Dispõe sobre os Planos de Benefícios da Previdência Social e dá outras providências**. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/18213cons.htm. Acesso em: 25 set. 2021.

BRITAIN, Sarah; BÖHNING, Dankmar. Estimators in capture-recapture studies with two sources. **AStA Advances in Statistical Analysis**, v. 93, n. 1, p. 23-47, 2009. DOI: 10.1007/s10182-008-0085-y. Disponível em: <https://doi.org/10.1007/s10182-008-0085-y>.

CIARAPICA, F. E.; GIACCHETTA, G. Classification and prediction of occupational injury risk using soft computing techniques: An Italian study. **Safety science**, v. 47, n. 1, p. 36-49, 2009. DOI: 10.1016/j.ssci.2008.01.006. Disponível em: <https://doi.org/10.1016/j.ssci.2008.01.006>.

COELI, Cláudia Medina; VERAS, Renato Peixoto; COUTINHO, Evandro da Silva Freire. Metodologia de captura-recaptura: uma opção para a vigilância das doenças não transmissíveis na população idosa. **Cadernos de Saúde Pública**, v. 16, p. 1071-1082, 2000. DOI: 10.1590/S0102-311X2000000400025. Disponível em: <https://doi.org/10.1590/S0102-311X2000000400025>.

CORREA, Paulo Roberto Lopes; ASSUNÇÃO, Ada Ávila. A subnotificação de mortes por acidentes de trabalho: estudo de três bancos de dados. **Epidemiologia e Serviços de Saúde**, v. 12, n. 4, p. 203-212, 2003. DOI: 10.5123/S1679-49742003000400004. Disponível em: <http://dx.doi.org/10.5123/S1679-49742003000400004>.

DA SILVA, Sabrina Letícia Couto; AMARAL, Fernando Gonçalves. Critical factors of success and barriers to the implementation of occupational health and safety management systems: A systematic review of literature. **Safety science**, v. 117, p. 123-132, 2019. DOI: 10.1016/j.ssci.2019.03.026. Disponível em: <https://doi.org/10.1016/j.ssci.2019.03.026>.

DATASUS, SIM-Sistema de Informações de Mortalidade, 2020. Disponível em: <<http://www2.datasus.gov.br/DATASUS/index.php?area=060701>>, acessado em 28 de junho 2020.

DEMIR, Pelin; GUL, Muhammet; GUNERI, Ali Fuat. Evaluating occupational health and safety service quality by SERVQUAL: a field survey study. **Total Quality Management & Business Excellence**, v. 31, n. 5-6, p. 524-541, 2020. DOI: 10.1080/14783363.2018.1433029. Disponível em: <https://doi.org/10.1080/14783363.2018.1433029>.

DIAS FERREIRA, Milene et al. Biological accidents underreporting by nursing professionals of a university hospital. **Cienc. enferm.**, Concepción, v. 21, n. 2, p. 21-29, Aug. 2015. DOI: 10.4067/S0717-95532015000200003. Disponível em: <http://dx.doi.org/10.4067/S0717-95532015000200003>.

FAN, Z. Joyce et al. Underreporting of work-related injury or illness to workers' compensation: individual and industry factors. **Journal of Occupational and Environmental Medicine**, v. 48, n. 9, p. 914-922, 2006. DOI: 10.1097/01.jom.0000226253.54138.1e. Disponível em: <https://doi.org/10.1097/01.jom.0000226253.54138.1e>.

FAYYAD, Usama; PIATETSKY-SHAPIRO, Gregory; SMYTH, Padhraic. The KDD process for extracting useful knowledge from volumes of data. **Communications of the ACM**, v. 39, n. 11, p. 27-34, 1996. DOI: 10.1145/240455.240464. Disponível em: <https://doi.org/10.1145/240455.240464>.

FELDMAN, Justin M. et al. Quantifying underreporting of law-enforcement-related deaths in United States vital statistics and news-media-based data sources: A capture-recapture analysis. **PLoS medicine**, v. 14, n. 10, 2017. DOI: 10.1371/journal.pmed.1002399. Disponível em: <https://doi.org/10.1371/journal.pmed.1002399>.

GALIZZI, Monica et al. Injured workers' underreporting in the health care industry: an analysis using quantitative, qualitative, and observational data. **Industrial Relations: A Journal of Economy and Society**, v. 49, n. 1, p. 22-43, 2010. DOI: 10.1111/j.1468-232x.2009.00585.x. Disponível em: <https://doi.org/10.1111/j.1468-232x.2009.00585.x>.

GERHARDT, Tatiana Engel; SILVEIRA, Denise Tolfo. Métodos de pesquisa. Plageder, 2009.

GIL, Antonio Carlos et al. Como elaborar projetos de pesquisa. São Paulo: Atlas, 2002.

GIRGIN, S.; KRAUSMANN, E. Historical analysis of US onshore hazardous liquid pipeline accidents triggered by natural hazards. **Journal of Loss Prevention in the Process Industries**, v. 40, p. 578-590, 2016. DOI: 10.1016/j.jlp.2016.02.008. Disponível em: <https://doi.org/10.1016/j.jlp.2016.02.008>.

GUL, Muhammet; AK, M. Fatih; GUNERI, Ali Fuat. Occupational health and safety risk assessment in hospitals: A case study using two-stage fuzzy multi-criteria approach. **Human and Ecological Risk Assessment: An International Journal**, v. 23, n. 2, p. 187-202, 2017. DOI: 10.1080/10807039.2016.1234363. Disponível em: <https://doi.org/10.1080/10807039.2016.1234363>.

HAJAKBARI, Maryam Sadat; MINAEI-BIDGOLI, Behrouz. A new scoring system for assessing the risk of occupational accidents: A case study using data mining techniques with Iran's Ministry of Labor data. **Journal of Loss Prevention in the Process Industries**, v. 32, p. 443-453, 2014. DOI: 10.1016/j.jlp.2014.10.013. Disponível em: <https://doi.org/10.1016/j.jlp.2014.10.013>.

HASSEL, Martin; ASBJØRNSLETT, Bjørn Egil; HOLE, Lars Petter. Underreporting of maritime accidents to vessel accident databases. **Accident Analysis & Prevention**, v. 43, n. 6, p. 2053-2063, 2011. DOI: . Disponível em: <https://doi.org/10.1016/j.aap.2011.05.027>.

ILO, 2018. Disponível em: https://www.ilo.org/wcmsp5/groups/public/---europe/---ro-geneva/---ilo-lisbon/documents/publication/wcms_726762.pdf. Acesso em 25 set. 2021.

INSS, Comunicação de Acidente de Trabalho – CAT, 2020, disponível em: <https://www.inss.gov.br/servicos-do-inss/comunicacao-de-acidente-de-trabalho-cat/>, acessado em 25 de junho de 2020.

ISO 45001, Disponível em: Occupational health and safety management systems — Requirements with guidance for use, 2020. <https://www.iso.org/standard/63787.html>, acessado em 28 de junho de 2020.

KAKHKI, Fatemeh Davoudi; FREEMAN, Steven A.; MOSHER, Gretchen A. Evaluating machine learning performance in predicting injury severity in agribusiness industries. **Safety science**, v. 117, p. 257-262, 2019. DOI: 10.1016/j.ssci.2019.04.026. Disponível em: <https://doi.org/10.1016/j.ssci.2019.04.026>.

KHOSROWABADI, Naghmeh; GHOSI, Rouzbeh; MAKUI, Ahmad. Decision Support Approach on Occupational Safety Using Data Mining. **International Journal of Industrial Engineering & Production Research**, v. 30, n. 2, p. 149-164, 2019. DOI: 10.22068/ijiepr.30.2.149. Disponível em: <http://dx.doi.org/10.22068/ijiepr.30.2.149>.

KODAPANAKKAL, Rabia I. et al. Self-interest and data protection drive the adoption and moral acceptability of big data technologies: A conjoint analysis approach. **Computers in Human**

Behavior, v. 108, p. 106303, 2020. DOI: 10.1016/j.chb.2020.106303. Disponível em: <https://doi.org/10.1016/j.chb.2020.106303>.

LIAO, Chia-Wen; PERNG, Yeng-Horng. Data mining for occupational injuries in the Taiwan construction industry. **Safety science**, v. 46, n. 7, p. 1091-1102, 2008. DOI: 10.1016/j.ssci.2007.04.007. Disponível em: <https://doi.org/10.1016/j.ssci.2007.04.007>.

MADSEN, Christian Uhrenholdt et al. Making occupational health and safety management systems ‘work’: A realist review of the OHSAS 18001 standard. **Safety Science**, v. 129, p. 104843, 2020. DOI: 10.1016/j.ssci.2020.104843. Disponível em: <https://doi.org/10.1016/j.ssci.2020.104843>.

MAHMOOD, Ibrahim Nasir; ALIEDANE, Hussein Ali Khuodhur; ABUZARAIDA, Mustafa Ali. Applications of Data Mining in Mitigating Fire Accidents Based on Association Rules. **International Journal of Interactive Mobile Technologies**, v. 15, n. 12, p. 159, 2021. DOI: 10.3991/ijim.v15i12.22687. Disponível em: <https://doi.org/10.3991/ijim.v15i12.22687>.

MARCZYK, Geoffrey; DEMATTEO, David. Essentials of research design and methodology. John Wiley & Sons, 2005.

NAGAO, Miki et al. Accidental exposures to blood and body fluid in the operation room and the issue of underreporting. **American Journal of Infection Control**, v. 37, n. 7, p. 541-544, 2009. DOI: 10.1016/j.ajic.2009.01.009. Disponível em: <https://doi.org/10.1016/j.ajic.2009.01.009>.

OLIVEIRA, Catarina Isabel Figueiredo. **Método de captura e recaptura para a estimação da abundância de uma população: aplicação da metodologia Bootstrap**. 145 f. Dissertação (Mestrado em Probabilidade e Estatística) - Universidade de Lisboa, Faculdade de Ciências, [S. l.], 2007. Disponível em: <https://repositorio.ul.pt/handle/10451/1232>.

PAPADOPOULOS, Gerasimos et al. Occupational and public health and safety in a changing work environment: An integrated approach for risk assessment and prevention. **Safety science**, v. 48, n. 8, p. 943-949, 2010. DOI: 10.1016/j.ssci.2009.11.002. Disponível em: <https://doi.org/10.1016/j.ssci.2009.11.002>.

POULIAKAS, Konstantinos; THEODOSSIOU, Ioannis. The economics of health and safety at work: an interdisciplinary review of the theory and policy. **Journal of Economic Surveys**, v. 27, n. 1, p. 167-208, 2013. DOI: 10.1111/j.1467-6419.2011.00699.x. Disponível em: <https://doi.org/10.1111/j.1467-6419.2011.00699.x>.

PROBST, T. M.; ESTRADA, A. X. Accident under-reporting among employees: Testing the moderating influence of psychological safety climate and supervisor enforcement of safety practices. **Accident Analysis and Prevention**, v. 42, n. 5, p. 1438–1444, 1 set. 2010. DOI: 10.1016/j.aap.2009.06.027. Disponível em: <https://doi.org/10.1016/j.aap.2009.06.027>.

PROBST, Tahira M. et al. Moderating effects of contingent work on the relationship between job insecurity and employee safety. **Safety science**, v. 106, p. 285-293, 2018. DOI: 10.1037/a0039195. Disponível em: <https://doi.org/10.1037/a0039195>.

PROBST, Tahira M. Organizational safety climate and supervisor safety enforcement: Multilevel explorations of the causes of accident underreporting. **Journal of Applied Psychology**, v. 100, n. 6, p. 1899, 2015. DOI: 10.1037/a0039195. Disponível em: <https://doi.org/10.1037/a0039195>.

PROBST, Tahira M.; GRASO, Maja. Pressure to produce= pressure to reduce accident reporting?. **Accident Analysis & Prevention**, v. 59, p. 580-587, 2013. DOI: 10.1016/j.aap.2013.07.020. Disponível em: <https://doi.org/10.1016/j.aap.2013.07.020>.

PROBST, Tahira M.; PETITTA, Laura; BARBARANELLI, Claudio. Comparing recall vs. recognition measures of accident under-reporting: A two-country examination. **Accident Analysis & Prevention**, v. 106, p. 1-9, 2017. DOI: 10.1016/j.aap.2017.05.006. Disponível em: <https://doi.org/10.1016/j.aap.2017.05.006>.

PRODANOV, Cleber Cristiano; DE FREITAS, Ernani Cesar. **Metodologia do trabalho científico: métodos e técnicas da pesquisa e do trabalho acadêmico-2ª Edição**. Editora Feevale, 2013.

RAMOS, Delfina; AFONSO, Paulo; RODRIGUES, Matilde A. Integrated management systems as a key facilitator of occupational health and safety risk management: A case study in a medium sized waste management firm. **Journal of Cleaner Production**, p. 121346, 2020. DOI: 10.1016/j.jclepro.2020.121346. Disponível em: <https://doi.org/10.1016/j.jclepro.2020.121346>.

REIS, Ruy Freitas et al. Characterization of the COVID-19 pandemic and the impact of uncertainties, mitigation strategies, and underreporting of cases in South Korea, Italy, and Brazil. **Chaos, Solitons & Fractals**, p. 109888, 2020. DOI: 10.1016/j.chaos.2020.109888. Disponível em: <https://doi.org/10.1016/j.chaos.2020.109888>.

RIVERA, Adovich Sarmiento; LAM, Hilton Yu. Applying the capture-recapture method to estimate road traffic deaths and injuries in three non-contiguous cities in the Philippines. **International journal of injury control and safety promotion**, v. 26, n. 1, p. 16-20, 2019. DOI: 10.1080/17457300.2018.1473447. Disponível em: <https://doi.org/10.1080/17457300.2018.1473447>.

RODRIGUES, Alana Barbosa; SANTANA, Vilma Sousa. Acidentes de trabalho fatais em Palmas, Tocantins, Brasil: oportunidades perdidas de informação. **Revista Brasileira de Saúde Ocupacional**, v. 44, 2019. DOI: 10.1590/2317-6369000017817. Disponível em: <https://doi.org/10.1590/2317-6369000017817>.

SÁNCHEZ, A. Suárez et al. Prediction of work-related accidents according to working conditions using support vector machines. **Applied Mathematics and Computation**, v. 218, n. 7, p. 3539-3552, 2011. DOI: 10.1016/j.amc.2011.08.100. Disponível em: <http://dx.doi.org/10.1016/j.amc.2011.08.100>.

SANGO, Hammadoum A. et al. Mortality and morbidity of urban road traffic crashes in Africa: capture-recapture estimates in Bamako, Mali, 2012. **PloS one**, v. 11, n. 2, 2016. DOI: 10.1371/journal.pone.0149070. Disponível em: <https://doi.org/10.1371/journal.pone.0149070>.

SARKAR, Sobhan et al. Application of optimized machine learning techniques for prediction of occupational accidents. **Computers & Operations Research**, v. 106, p. 210-224, 2019. DOI: 10.1016/j.cor.2018.02.021. Disponível em: <https://doi.org/10.1016/j.cor.2018.02.021>.

SHANNON, Harry S.; LOWE, Graham S. How many injured workers do not file claims for workers' compensation benefits?. **American journal of industrial medicine**, v. 42, n. 6, p. 467-473, 2002. DOI: 10.1002/ajim.10142. Disponível em: <https://doi.org/10.1002/ajim.10142>.

SHORT, Jack; CAULFIELD, Brian. Record linkage for road traffic injuries in Ireland using police hospital and injury claims data. **Journal of safety research**, v. 58, p. 1-14, 2016. DOI: 10.1016/j.jsr.2016.05.002. Disponível em: <https://doi.org/10.1016/j.jsr.2016.05.002>.

SINAN, 2020. Disponível em: <<http://portalsinan.saude.gov.br/#:~:text=O%20Sistema%20de%20Informa%C3%A7%C3%A3o%20de,I%2C%20mas%20%C3%A9%20facultado%20a>>, acessado em 28 de junho 2020.

SINGH, Arpit; MISRA, Subhas Chandra. A dominance based rough set analysis for investigating employee perception of safety at workplace and safety compliance. **Safety science**, v. 127, p. 104702, 2020. DOI: 10.1016/j.ssci.2020.104702. Disponível em: <https://doi.org/10.1016/j.ssci.2020.104702>.

SKŁAD, Anna. Assessing the impact of processes on the Occupational Safety and Health Management System's effectiveness using the fuzzy cognitive maps approach. **Safety science**, v. 117, p. 71-80, 2019. DOI: 10.1016/j.ssci.2019.03.021. Disponível em: <https://doi.org/10.1016/j.ssci.2019.03.021>.

TEUFER, Birgit et al. Evidence-based occupational health and safety interventions: a comprehensive overview of reviews. **BMJ open**, v. 9, n. 12, 2019. DOI: 10.1136/bmjopen-2019-032528. Disponível em: <http://dx.doi.org/10.1136/bmjopen-2019-032528>.

TUCKER, Sean et al. Work-related injury underreporting among young workers: prevalence, gender differences, and explanations for underreporting. **Journal of safety research**, v. 50, p. 67-73, 2014. DOI: 10.1016/j.jsr.2014.04.001. Disponível em: <https://doi.org/10.1016/j.jsr.2014.04.001>.

VARIANOU-MIKELLIDOU, Cleo et al. Work-related factors and individual characteristics affecting work ability of different age groups. **Safety Science**, v. 128, p. 104755, 2020. DOI: 10.1016/j.ssci.2020.104755. Disponível em: <https://doi.org/10.1016/j.ssci.2020.104755>.

VOIDE, Cathy et al. Underreporting of needlestick and sharps injuries among healthcare workers in a Swiss University Hospital. **Swiss medical weekly**, v. 142, n. w13523, p. 1-7, 2012. DOI: 10.4414/smw.2012.13523. Disponível em: <https://doi.org/10.4414/smw.2012.13523>.

WUELLNER, Sara E.; BONAUTO, David K. Exploring the relationship between employer recordkeeping and underreporting in the BLS Survey of Occupational Injuries and Illnesses. **American journal of industrial medicine**, v. 57, n. 10, p. 1133-1143, 2014. DOI: 10.1002/ajim.22350. Disponível em: <https://doi.org/10.1002/ajim.22350>.

ZADOW, Amy Jane et al. Psychosocial safety climate, emotional exhaustion, and work injuries in healthcare workplaces. **Stress and Health**, v. 33, n. 5, p. 558-569, 2017. DOI: 10.1002/smi.2740. Disponível em: <https://doi.org/10.1002/smi.2740>.

ZHU, Yuan et al. SAVE-T: Safety Analysis Visualization and Evaluation Tool. **Journal of Advanced Transportation**, v. 2021, 2021. DOI: 10.1155/2021/5545117. Disponível em: <https://doi.org/10.1155/2021/5545117>.