



MINISTÉRIO DA SAÚDE
INSTITUTO NACIONAL DE CARDIOLOGIA
COORDENAÇÃO DE ENSINO E PESQUISA
PROGRAMA DE PÓS-GRADUAÇÃO
MESTRADO PROFISSIONAL EM AVALIAÇÃO DE TECNOLOGIAS EM SAÚDE

SOLANGE BORGES DA SILVA CORDEIRO

EXTRAPOLAÇÃO DE DADOS COM CURVAS DE SOBREVIDA: uma
proposta metodológica para dados agregados

RIO DE JANEIRO

2021

SOLANGE BORGES DA SILVA CORDEIRO

EXTRAPOLAÇÃO DE DADOS COM CURVAS DE SOBREVIDA: uma
proposta metodológica para dados agregados

Dissertação apresentada como
requisito à obtenção do título de
Mestre em Avaliação de Tecnologias
em Saúde (ATS), ao Programa de
Pós-Graduação do Instituto Nacional
de Cardiologia, Mestrado Profissional
em Avaliação de Tecnologias em
Saúde (ATS)

Orientador: Ivan Ricardo Zimmermann

RIO DE JANEIRO

2021

FICHA CATALOGRÁFICA

C794e Codeiro, Solange Borges da Silva.

Extrapolação de dados com curvas de sobrevida:
uma proposta metodológica para dados agregados /
Solange Borges da Silva Cordeiro. – Rio de Janeiro, 2021.

90 f.

Dissertação (Mestrado Profissional em
Avaliação de Tecnologias em Saúde) Instituto Nacional
de Cardiologia – INC

SOLANGE BORGES DA SILVA CORDEIRO

EXTRAPOLAÇÃO DE DADOS COM CURVAS DE SOBREVIVÊNCIA: uma proposta metodológica para dados agregados

Dissertação apresentada como requisito à obtenção do título de Mestre em Avaliação de Tecnologias em Saúde (ATS), ao Programa de Pós-Graduação do Instituto Nacional de Cardiologia, Mestrado Profissional em Avaliação de Tecnologias em Saúde (ATS)

Aprovada em:

Banca Examinadora:

Prof. Dr. Ivan Zimmermann
Orientador
Universidade de Brasília

Prof. Dr. Carlos Alberto da Silva Magliano
Membro interno
Instituto Nacional de Cardiologia

Prof. Dr. Bernardo Rangel Tura
Membro interno
Instituto Nacional de Cardiologia

Prof. Dr^a. Patrícia Klarmann Ziegelmann
Membro externo
Universidade Federal do Rio Grande do Sul

Prof. Dr^a. Marisa Santos
Membro interno (suplente)
Instituto Nacional de Cardiologia

Prof. Dr. Ricardo Fernandes
Membro externo (suplente)
Instituto Nacional do Câncer

Este trabalho é dedicado a Deus, por ter me capacitado e estado comigo em todos os instantes de minha vida. Dedico também aos meus amados pais, Ione Alves da Silva e José Borges da Silva por ser meu exemplo de luta que sempre me encorajaram. Dedico aos meus amados, Fábio da Silva Cordeiro e meu filho Kalel Borges Cordeiro por serem meu porto seguro, amigos e companheiros inseparáveis sempre pacientes e compreensivos. E a minha amiga Iane Coutinho, que me ajudou e incentivou a entrar neste grande desafio. Por fim, dedico aos meus irmãos, cunhadas e meus sobrinhos, por todo apoio e ajuda para que esse momento se tornasse possível. Amo muito vocês.

AGRADECIMENTOS

Gostaria de agradecer a todos aqueles que de uma forma ou de outra estiveram presentes em minha vida me apoiando durante todos esses anos em especial aos meus pais, José Borges e Ione, dos quais tenho imenso orgulho, admiração e respeito. Gostaria de agradecer a meu marido, Fábio, o qual foi o grande incentivador de toda a minha jornada no mestrado e na vida. Ao meu amado filho Kalel, por me trazer alegria nos momentos difíceis. Família, amo vocês, pois foram indispensáveis para que essa conquista se realizasse.

Gostaria de aproveitar e agradecer a todos os amigos que fiz no mestrado, por terem me dado um importante auxílio durante o curso, para que eu pudesse superar mais uma difícil batalha.

Gostaria de agradecer especialmente ao professor Dr. Ivan Ricardo Zimmerman por toda paciência, disponibilidade, motivação e qualidade didática no auxílio deste trabalho.

Agradeço o empenho e a força de vontade da professora Dr.^a Marisa da Silva Santos, que muito se esforça para que este curso faça a diferença na vida das pessoas que por aqui passam, com certeza meu olhar será com o objetivo de enxergar além do que me for apresentado e pensar sempre no bem coletivo.

E a todos os professores que conheci aqui no INC que me proporcionaram a oportunidade do conhecimento em especial aos professores Dr. Bernardo Rangel Tura que me fez olhar a matemática de forma desgarrada de suas fórmulas e regras, Márcia Gisele Costa e Kátia Marie Simões e Senna pelo zelo e qualidade nos trabalhos, sendo incansáveis na melhoria dos mesmos.

“Combati o bom combate, completei a corrida, perseverei na fé!”

2 Timóteo 4:7

RESUMO

Introdução: Muitas pesquisas clínicas possuem curto tempo de seguimento dos participantes, tornando-se necessária a adoção de métodos de extrapolação de dados nos modelos de avaliação econômica. Apesar disso, as propostas metodológicas disponíveis trabalham com os dados em sua forma individualizada, cenário nem sempre factível ao cotidiano da Avaliação de Tecnologias em Saúde (ATS). Assim, o objetivo deste estudo foi reproduzir um método para extrapolação com curvas de sobrevida aplicável a dados agregados, que são dados de sobrevida que não estão disponíveis a nível de indivíduos (Individual Patient data – IPD). **Métodos:** Inicialmente, revisou-se as diretrizes e recomendações sobre extrapolação de dados e em seguida, as aplicaram em um estudo de caso. A revisão rápida consistiu em busca no Medline (via Pubmed), nas 10 principais agências de ATS e comunicações enviadas a rede internacional de agências de ATS, *The International Network of Agencies for Health Technology Assessment* (INAHTA). Foram incluídas diretrizes e demais documentos com recomendações explícitas relacionadas ao tema da extrapolação de dados de sobrevida na avaliação econômica em saúde, sem restrição de período, nos idiomas inglês, espanhol e português. Os dados foram extraídos em formulário padrão e realizada síntese qualitativa. O estudo de caso consistiu nas aplicações de modelos não paramétricos e paramétricos de análise de sobrevivência em duas terapias (anastrozol e fulvestranto) no contexto do câncer de mama metastático com dados de 10 anos de seguimento nos desfechos de sobrevida livre de progressão (SLP) e sobrevida global (SG) para dois braços de tratamento. Para o propósito do trabalho, buscou-se a aplicação do ajuste de curvas de sobrevida com o método de mínimos quadrados não lineares aos dados agregados. Os autores do estudo primário foram contactados sobre a possibilidade de compartilhamento de dados individualizados para a eventual comparação e validação do método. **Resultados:** Foram encontrados 306 documentos, três foram incluídos no processo de seleção. As principais distribuições recomendadas e elegidas para o estudo de caso foram: Exponencial, Weibull, Log-normal, Log-logística, Gama generalizada e Gompertz. Na verificação da aplicação no estudo de caso nos dados dos dois braços de tratamentos, as distribuições Weibull e gama generalizada foram as que melhor se ajustaram aos dados de SG, seja na abordagem individualizada quanto na agregada, segundo critérios estatísticos e de inspeção visual. Para a SLP foram as curvas log-logística e log-normal para o Anastrozol. No caso do tratamento com fulvestranto,

foram as curvas log-normal e gama generalizada para os dados individualizados e Gompertz e gama generalizada nos dados agregados. **Conclusão:** Foi possível mostrar a factibilidade do ajuste de dados com curvas de sobrevida a partir de dados agregados com o objetivo de extrapolação de dados em modelos econômicos. Apesar das divergências, ao se compararem os modelos sugeridos por ambos os métodos, observam-se que seria pouco provável a existência de impacto relevantes para a tomada de decisão. Tal método pode ser útil nas situações de impossibilidade de acesso aos dados individualizados.

Palavras-chave: Análise de Sobrevida; Avaliação de Custo-Efetividade; Avaliação da Tecnologia Biomédica; Guias como Assunto; Oncologia

ABSTRACT

Introduction: Many clinical researches have a short follow-up time of the participants, making it necessary to adopt methods for extrapolating data in the economic evaluation models. Despite this, the methodological approaches available only handle with individualized patient data, a scenario not always feasible in Health Technology Assessment (HTA). Thus, the objective of this study was to reproduce a method for extrapolation with survival curves applicable to aggregated data, which are survival data that are not available at the individual level (Individual Patient data - IPD).

Methods: Initially, the guidelines and recommendations on data extrapolation were reviewed and then these recommendations were applied in a case study. The rapid review consisted of searching Medline (via Pubmed), the main HTA agencies and the international HTA network members (INAHTA). Methodological guidelines and other documents with explicit recommendations related to the topic of extrapolation of survival data in health economic evaluation were included, without publication date restriction, in English, Spanish and Portuguese languages. The data were extracted in a standard form and a qualitative synthesis was carried out. The case study consisted of the application of non-parametric and parametric models of survival analysis in two therapies (anastrozole and fulvestrant) in the context of metastatic breast cancer with 10-year follow-up data on progression-free survival (SLP) and overall survival (SG) outcomes. For this purpose, we have applied the adjustment of survival curves to the aggregated data with nonlinear least squares methods. The authors of the primary study were contacted about the possibility of sharing individual data for comparison and validation of the method. **Results:** 306 documents were retrieved, of which only three were included at the end of the selection process. The main distributions recommended and chosen for the case study were: Exponential, Weibull, Log-normal, Log-logistic, Generalized gamma and Gompertz. In verifying the application in the case study in the data of the two treatment arms, the Weibull and generalized gamma distributions were the ones that best fit the SG data, both in the individualized and in the aggregate approach, according to statistical and visual inspection criteria. For SLP were the log-logistic and log-normal curves for Anastrozole. In the case of treatment with fulvestrant, it was the log-normal and generalized gamma curves for the individualized data and Gompertz and generalized gamma for the aggregated data.

Conclusion: It was possible to show the feasibility of adjusting data with survival curves from aggregated data in order to extrapolate data in economic models. Despite some divergences, when comparing the models suggested by both methods, it was observed that the existence of relevant impacts for decision-making would be unlikely. Such a method can be useful in situations where it is impossible to access individualized data.

Keywords: Survival Analysis; Cost-Effectiveness Evaluation; Technology Assessment, Biomedical; Guidelines as Topic; Medical Oncology.

LISTA DE ILUSTRAÇÕES

Figura 1 – Fluxo de incorporação de tecnologias no Brasil.	20
Figura 2 – Mecanismo de censura à direita	23
Figura 3 – Divisão dos grupos, para uso dos medicamentos	29
Figura 4 – Curvas de Kaplan-Meier apresentadas no estudo.....	31
Figura 5 – Curva de sobrevivência de Kaplan-Meier dos dados individualizados versus dados agregados	32
Figura 6 – Curva teórica da função de sobrevivência	33
Figura 7 - Curva teórica da função de risco	34
Figura 8 – Fluxograma de seleção das publicações	40
Figura 9 – Processo para escolha do modelo de sobrevivência.....	44
Figura 10 – Curva de sobrevivência de Kaplan-Meier versus curvas de sobrevivência estimadas propostas para os dados individualizados – tratamento Anastrozol, sobrevida global e livre de progressão	51
Figura 11 – Curva de sobrevivência de Kaplan-Meier versus curvas de sobrevivência estimadas propostas para os dados individualizados – tratamento Anastrozol+Fulvestranto, Sobrevida global e livre de progressão	52
Figura 12 – Curva de sobrevivência de Kaplan-Meier versus curvas de sobrevivência estimadas propostas – tratamento Anastrozol, Sobrevida global e Livre de progressão.....	56
Figura 13 – Curva de sobrevivência de Kaplan-Meier versus curvas de sobrevivência estimadas propostas – tratamento Anastrozol+Fulvestranto, Sobrevida global e livre de progressão	57
Gráfico 1 – Curvas de Kaplan-Meier obtidas dos dados individualizados para os dois braços de tratamento.....	47
Gráfico 2 – Curvas obtidas das figuras digitalizadas para os dois braços de tratamento	48
Gráfico 3 – Curvas de Kaplan-Meier obtidas dos dados individualizados para o Anastrozol vs distribuições estimadas.....	49

Gráfico 4 – Curvas de Kaplan-Meier obtidas dos dados individualizados para o Anastrozol+Fulvestranto vs distribuições estimadas.....	50
Gráfico 5 – Curvas da Kaplan-Meier obtidas dos dados agregados para o Anastrozol vs distribuições estimadas – sobrevida global e livre de progressão	54
Gráfico 6 – Curvas da Kaplan-Meier obtidas dos dados agregados para o Anastrozol+Fulvestranto vs distribuições estimadas – sobrevida global e livre de progressão	55
Gráfico 7 – Comparação das curvas com dados individuais vs agregados, sobrevida global e sobrevida livre de progressão.....	59
Quadro 1 – Estratégia de busca e resultados selecionados	27
Quadro 2 – Principais Agências de ATS	28
Quadro 3 – Síntese dos resultados obtidos.....	41

LISTA DE TABELAS

Tabela 1: Estatísticas do critério de informação AIC e BIC, sobrevida global e livre de progressão, para os tratamentos Anastrozol e Anastrozol+Fulvestranto, para os dados individualizados.....	53
Tabela 2: Estatísticas do critério de informação AIC e BIC, sobrevida global e livre de progressão, para os tratamentos Anastrozol e Anastrozol+Fulvestranto, para os dados agregados	58

LISTA DE ABREVIATURAS E SIGLAS

AIC	Akaike's Information Criterion
ASPS	Ações e Serviços Públicos de Saúde
ATS	Avaliação de Tecnologia em Saúde
BIC	Bayesian Information Criterion
CITEC	Comissão de Incorporação de Tecnologias
CONITEC	Comissão Nacional de Incorporação de Tecnologias no SUS
DCNT	Doenças Crônicas Não Transmissíveis
DECIT	Departamento de Ciência e Tecnologia
EC	Emenda Constitucional
INAHTA	International Network of Agencies for Health Technology Assessment
IPEA	Instituto de Pesquisa Econômica Aplicada
IPD	Individual Patient Data
MS	Ministério da Saúde
NICE	National Institute for Health and Clinical Excellence
NLS	Nonlinear Least Squares
PBAC	Pharmaceutical Benefits Advisory Committee
REBRATS	Rede Brasileira de Avaliação de Tecnologias em Saúde
SE	Secretaria Executiva
SUS	Sistema Único de Saúde

SUMÁRIO

1 INTRODUÇÃO	18
1.1 A Avaliação de Tecnologias em Saúde (ATS)	18
1.1.2 A incorporação de tecnologias no Sistema Único de Saúde (SUS)	19
1.2 Avaliação econômica em saúde	20
1.2.1 Lidando com o horizonte temporal	21
1.3 Análise de Sobrevivência	22
1.3.1 Algumas definições	22
1.3.2 Modelos de análise de sobrevivência	24
1.4 A ATS na especialidade da oncologia	24
1.4.1 O câncer de mama no Brasil e no mundo	25
2 OBJETIVO GERAL	26
2.1 Objetivos específicos	26
3 MÉTODOS	27
3.1 Revisão das diretrizes disponíveis	27
3.1.1 Tipo de estudo	27
3.1.2 Estratégia de busca	27
3.1.3 Critérios de seleção	28
3.1.4 Extração	29
3.1.5 Síntese	29
3.2 Aplicação das diretrizes disponíveis	29
3.2.1 O estudo de caso no câncer de mama metastático	29
3.2.2 A digitalização de imagens	30
3.2.2.1 Funções de sobrevivência	33
3.2.2.2 Função de risco	34
3.2.2.2 Distribuição Exponencial	34
3.2.2.3 Distribuição de Weibull	35
3.2.2.4 Distribuição Log-normal	36
3.2.2.5 Distribuição de Gompertz	36
3.2.2.6 Distribuição Log-logística	37
3.2.2.7 Distribuição Gama Generalizada	37
3.3.1 Ajuste dos modelos	38

4 RESULTADOS	40
4.1 Revisão Rápida das diretrizes	40
4.1.1 PBAC	41
4.1.2 NICE	42
4.1.3 CONITEC	45
4.2 Ajuste de curvas aos dados no câncer de mama	46
4.2.1 Dados de sobrevivência observada	46
4.2.2 Dados de sobrevivência ajustados	48
5 DISCUSSÃO	60
5.1 Sobre as diretrizes	60
5.2 Sobre as extrapolações	61
6 CONCLUSÕES	64
REFERÊNCIAS	65
APÊNDICES	69
ANEXO	87

1 INTRODUÇÃO

1.1 A Avaliação de Tecnologias em Saúde (ATS)

A mudança no perfil epidemiológico e estrutura etária da população das últimas décadas tem pressionado os Sistemas de Saúde[1]. As doenças infecciosas e parasitárias que faziam parte do cenário mundial nos últimos cem anos e no Brasil dividem o espaço com as doenças crônicas não-transmissíveis (DCNT) a partir de 1960[2], por sua vez, com custos crescentes. Como os recursos não são ilimitados, deve-se esperar dos gestores em saúde um direcionamento de qual é a tecnologia mais indicada, ou seja, que tenha um melhor balanço de riscos e benefícios para ser disponibilizada à população. Em se tratando dos gastos brasileiros em saúde, o cenário é ainda mais preocupante se comparado ao de países desenvolvidos. Segundo estudos do Instituto de Pesquisa Econômica Aplicada (IPEA), a participação federal vem decrescendo no compartilhamento total dos gastos com serviços públicos em saúde no período de 2003 a 2017 passando de 50,1% para 43,2%. Além do histórico na diminuição dos gastos federais, temos o teto do orçamento federal para o Sistema Único de Saúde (SUS) de 2017 a 2036 (segundo a EC nº 95 em dezembro de 2016, que implantou o chamado Novo Regime Fiscal. Tal medida congelou, em termos reais, as despesas primárias da União e estabeleceu novo regime para a contabilização das despesas). Porém, foram desconsideradas todas as vertentes que indicam uma necessidade no aumento dos recursos.

Da mesma forma, a profícua disponibilidade dos resultados de pesquisas científicas torna complexo o trabalho de delimitar quais tecnologias são de fato superiores em termos de efetividade e segurança e, portanto, úteis para a população a que se pretende disponibilizar. Devido ao cenário de complexidade de informação e recursos escassos na área da saúde, tornou-se necessária a implantação de mecanismos de suporte para a decisão dos gestores de saúde de como e onde serão ou não alocados os recursos. Com o objetivo de assessorar as discussões sobre custos e benefícios da implementação de tecnologias em sistemas de saúde, dissemina-se o campo da Avaliação de Tecnologias em Saúde (ATS), que pode ser compreendida como:

[...] o processo contínuo de análise e síntese dos benefícios para a saúde, das consequências econômicas e sociais do emprego das tecnologias, considerando os seguintes aspectos: segurança, acurácia, eficácia, efetividade, custos, custo-

efetividade e aspectos de equidade, impactos éticos, culturais e ambientais envolvidos na sua utilização.(BRASIL, 2010, p 17)[3].

Nesse contexto, tem-se a inda que a Gestão de Tecnologias em Saúde é a prática responsável pelo estabelecimento de padrões e práticas de incorporação de tecnologias, sendo definida como:

“[...] conjunto de atividades gestoras relacionado com os processos de avaliação, incorporação, difusão, gerenciamento da utilização e retirada de tecnologias do sistema de saúde[...]”.(BRASIL, 2010, p 10)[3].

Nesse contexto, foram criadas em vários países, incluindo o Brasil, instâncias de ATS como uma estratégia de sua institucionalização, buscando subsidiar os decisores com informações para a tomada de decisão na formulação e implementação das políticas em saúde [4].

1.1.2 A incorporação de tecnologias no Sistema Único de Saúde (SUS)

No Brasil, pode-se dizer que o processo de institucionalização da ATS no Sistema Único de Saúde (SUS) teve seu início no Ministério da Saúde (MS) com a criação do Departamento de Ciência e Tecnologia (DECIT) em 2000. Posteriormente, em dezembro de 2006, é então instituída a Comissão de Incorporação de Tecnologias do Ministério da Saúde (CITEC) por meio da portaria Nº 3.323/2006 [5]. Em 2008, é a vez da criação da Rede Brasileira de Avaliação de Tecnologias em Saúde (REBRATS), sendo instituída em 2011, de acordo com a portaria Nº 2.915/2011, com seu principal objetivo:

Art. 2º A REBRATS é uma rede de centros colaboradores e instituições de ensino e pesquisa no País voltada à geração e à síntese de evidências científicas no campo de Avaliação de Tecnologias em Saúde (ATS) no Brasil e no âmbito internacional.(BRASIL, 2011)[6]

Em 2011, com a alteração da lei nº 8080/1990 pela lei **12.401/2011** [7], é criada a Comissão Nacional de Incorporação de Tecnologias no Sistema Único de Saúde (Conitec), substituindo a antiga Comissão de Incorporação de Tecnologias (CITEC). Desde então, a Conitec é o órgão colegiado responsável por assessorar o Ministério

da Saúde (MS) nas decisões sobre incorporar, alterar ou excluir novas tecnologias em saúde no âmbito do SUS. Além de constituir ou alterar os Protocolos Clínicos e Diretrizes Terapêuticas (PCDT).

A partir deste marco, a incorporação de tecnologias em saúde no SUS segue um claro fluxo central de avaliação, ilustrado pela Figura 1:



Figura 1 – Fluxo de incorporação de tecnologias no Brasil.

Fonte: CONITEC, 2016[8].

Em suas diretrizes metodológicas, esclarecesse-se que as avaliações das tecnologias submetidas à Conitec devem considerar, necessariamente, as evidências científicas sobre a eficácia, a acurácia, a efetividade e a segurança, e ainda as análises econômicas comparativas dos benefícios e dos custos em relação às tecnologias já incorporadas. Tais orientações são consistentes com a Lei nº 12.401/2011, que estabelece a ATS como critério indispensável para a tomada de decisão sobre a incorporação tecnológica no SUS [9].

1.2 Avaliação econômica em saúde

Considerada um dos grandes pilares da ATS, a avaliação econômica é definida como uma análise técnica que identifica, mensura, valora e compara estratégias em relação aos seus custos e desfechos em saúde [10]. Ou seja, observa o potencial de

que a escolha de uma estratégia em detrimento da outra, não tem efeito só em relação a saúde, mas também sobre os recursos que serão alocados.

Pode-se delimitar as avaliações econômicas em saúde em quatro tipos fundamentais: custo-efetividade, custo-minimização, custo-utilidade e custo-benefício. Todas as quatro análises avaliam os custos em termos monetários, apresentando suas diferenças em relação aos desfechos ou consequências. A custo-minimização ou minimização de custos seria um caso particular de duas ou mais intervenções que tenham eficácia semelhante. Dado o pressuposto de equivalência, o foco é monetário, em que a escolha da tecnologia se dará pela que acarretar menores custos. Na análise de custo-efetividade, os desfechos são medidos em unidades clínicas, como por exemplo, anos de vida ganhos ou número de eventos indesejáveis que deixaram de ocorrer. Como resultado, tal análise nos fornece as possíveis razões de custo-efetividade incrementais. Uma limitação deste tipo de estudo seria apenas comparar intervenções que tenham o mesmo desfecho em saúde. Na custo-utilidade, os desfechos em saúde utilizam especificamente alguma unidade de utilidade, como os anos de vida ajustados por qualidade (do inglês, quality adjusted life years – QALY) que seria a combinação de benefícios em tempo de vida (sobrevida) e em qualidade de vida (preferências ou utilidade). Por fim, na análise de custo-benefício, tanto os custos quanto os benefícios em saúde são medidos em unidades monetárias. Nessa análise específica, calcula-se o valor do benefício financeiro líquido, sendo interessante a estratégia com um saldo positivo[9,10].

1.2.1 Lidando com o horizonte temporal

Ao lidar com modelos matemáticos, uma avaliação econômica precisa definir uma série de critérios, dentre eles, o horizonte temporal pelo qual o modelo irá considerar os custos e as consequências. Para tanto, é crucial que o modelo tenha como base o curso natural da condição e o provável impacto que a intervenção tenha sobre ele. Segundo a diretriz metodológica de avaliação econômica brasileira, as doenças crônicas devem considerar a expectativa de vida dos pacientes como horizonte temporal e nas doenças ou condições agudas, que não acarretem sequelas a longo prazo, é indicado que utilize o menor período que permita avaliar os desfechos relevantes [9].

Na maioria das intervenções, a efetiva adoção de um horizonte de tempo que contemple a expectativa de vida implicará a necessidade de extrapolação de dados

provenientes da pesquisa clínica. Isto é particularmente verdade quando as evidências são provenientes de ensaios clínicos randomizados, onde os custos e barreiras pragmáticas de implementação impedem que o seu tempo de seguimento seja longo o suficiente para abarcar toda a expectativa de vida de seus participantes. Nessas situações, os cálculos da expectativa de vida adotada no modelo podem exigir a adoção de extrapolação de dados. Recentemente, a adição de modelos com curvas de sobrevivência para que o horizonte do modelo considere a incerteza na extrapolação dos dados disponíveis tem sido frequentemente adotado [11].

1.3 Análise de Sobrevivência

1.3.1 Algumas definições

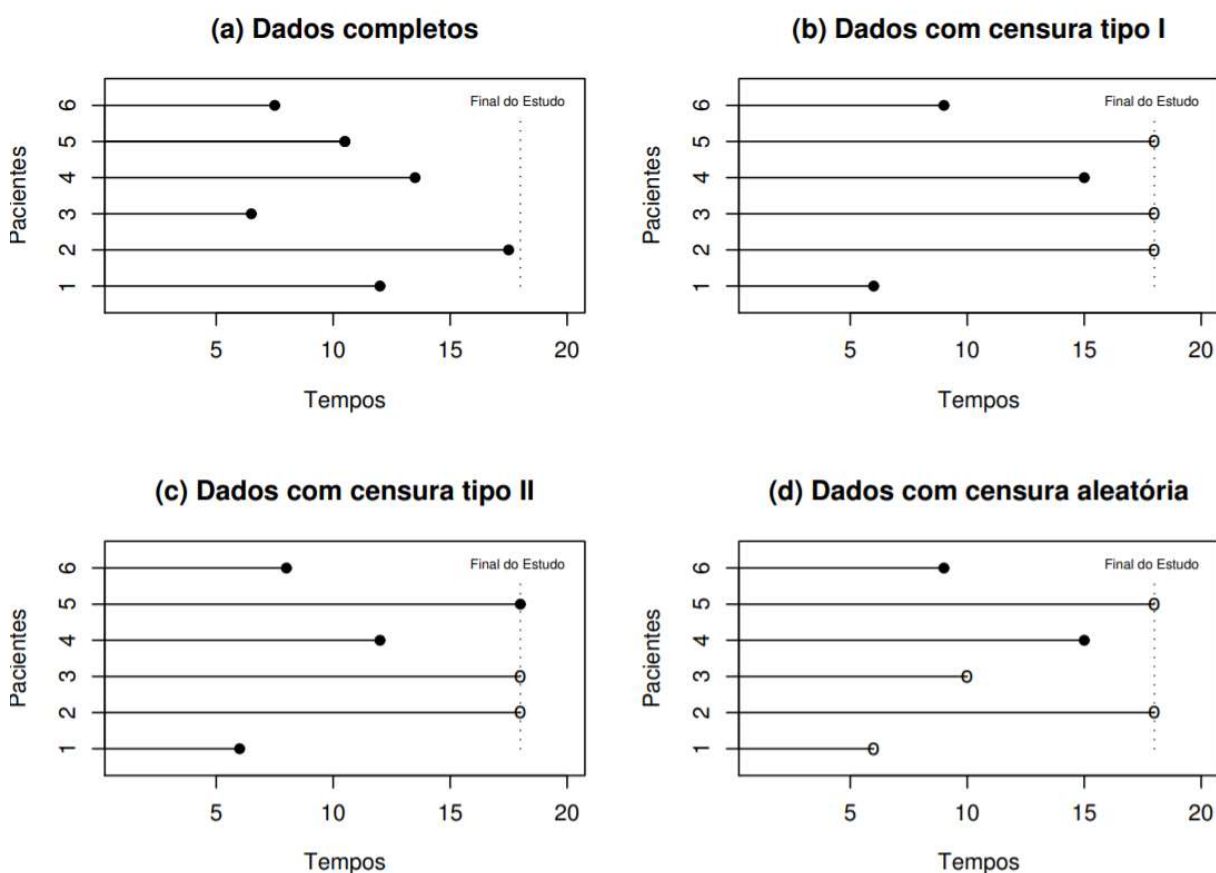
A análise de sobrevivência tem sido uma das áreas da estatística que mais cresceu a partir dos anos 80. Mostrando-se uma ferramenta importante na área da Medicina para estimar curvas de sobrevivência. Sendo tais curvas estimadas utilizadas para comparações entre dois ou mais tratamentos ao longo do tempo ou verificar quando se dará a ocorrência de um determinado desfecho de interesse, como óbito, cura, doença entre outros [12].

Na análise de sobrevivência, a variável resposta é o tempo até a ocorrência de um evento de interesse. O evento pode ser a morte ou a cura do paciente, a ocorrência ou a recorrência de uma doença, por exemplo. Este tempo até o evento se denomina tempo de falha, que é um termo igualmente muito usado no contexto de análise de confiabilidade [12]. Quando o acompanhamento de um paciente é interrompido por alguma eventualidade, temos a presença de censura, que é a observação parcial da resposta, esta é a principal característica da análise de sobrevivência. De maneira geral, as censuras podem ser à direita, à esquerda ou intervalar. A censura à direita ocorre à direita do tempo verificado, o tempo entre o começo do estudo e o evento de interesse é maior que o tempo observado. Existem três tipos de mecanismos de censura à direita:

- Censura do tipo I: ocorre quando o período de análise do estudo finda e ainda existem indivíduos que não apresentaram o evento de interesse, também chamada de censura administrativa;

- Censura do tipo II: quando o estudo é finalizado após a ocorrência de um número pré-estabelecido de evento de interesse, independente da possibilidade ou não de novas ocorrências;
- Censura do tipo aleatória: mais comum na prática, acontece quando o indivíduo é retirado da amostra do estudo sem ter ocorrido o evento de interesse, como por exemplo, se o indivíduo morrer por uma causa diferente da estudada. Os tipos de mecanismo de censura à direita estão ilustrados na Figura 2[13].

A Censura à esquerda acontece antes do tempo observado, ou seja, o indivíduo já sofreu o evento de interesse quando foi observado. Contudo, é desconhecido o momento da ocorrência. A censura intervalar acontece em estudos que possuem uma periodicidade para observação, ou seja, o evento ocorreu durante este período, não sabendo o momento exato da ocorrência. Mesmo havendo censura, os dados devem ser considerados no estudo, pois sua omissão pode levar a conclusões equivocadas. Neste trabalho os dados são censurados à direita.



○ Representa censura; ● Representa a falha

Figura 2 – Mecanismo de censura à direita

Fonte: Colosimo, 2006[12].

Outra característica em estudos de análise de sobrevivência, é o dado truncado, que é a perda total da observação, ou seja, são incluídos no estudo somente os indivíduos em que o evento de interesse tenha ocorrido dentro de um intervalo pré-definido [12,14].

1.3.2 Modelos de análise de sobrevivência

Os parâmetros dos modelos de análise de dados de sobrevivência podem ser estimados de forma não paramétrica, paramétrica ou semiparamétrica.

Nos modelos não paramétricos, as estimativas são realizadas sem atribuir uma distribuição de probabilidade conhecida para o tempo de sobrevivência, sendo os estimadores de função de sobrevivência ($S(t)$) e função de risco ($h(t)$) obtidos diretamente dos dados. Como exemplo frequente desta abordagem, cita-se o estimador de Kaplan-Meier (K-M). Já os modelos paramétricos supõem uma distribuição conhecida de probabilidade e seus parâmetros são estimados por meio de ajuste dos dados disponíveis. Como exemplo também frequente, cita-se a distribuição exponencial. Por fim, os modelos semiparamétricos permitem analisar o tempo até a ocorrência de um evento de interesse, ajustado por covariáveis. Neste caso, cita-se o exemplo do modelo de Cox [14].

1.4 A ATS na especialidade da oncologia

A oncologia é uma das principais especialidades demandadas por avaliação na Conitec. Em uma análise publicada da atividade da comissão, as principais demandas identificadas foram referentes a antineoplásicos e agentes imunomoduladores, correspondendo por mais de um terço dos medicamentos incorporados de 2012 a 2016[15]. Dentre os desfechos comumente relatados na pesquisa clínica da oncologia, destacam-se as estimativas de sobrevida-livre de progressão (SLP), que é o período de tempo durante o qual o tratamento está mantendo a doença controlada, ou seja, sem progressão e a sobrevida global (SG), sendo o tempo de vida após o início de um determinado tratamento [16].

Dada a limitação dos tempos de seguimento da pesquisa clínica e dos recentes avanços na modelagem, a adoção de técnicas de extrapolação de dados com curvas

de sobrevivência tem sido frequente nos modelos em oncologia. Apesar disso, orientações e recomendações sobre os critérios e métodos a serem adotados nesta extrapolação ainda são escassos. Segundo o instituto inglês de ATS NICE (do inglês, National Institute for Health and Care Excellence – NICE), no seu documento de suporte técnico (TSD14), que trata de análise de sobrevivência para avaliação econômica, foi verificado numa revisão de 45 estudos submetidos na área, 32 usaram métodos paramétricos para estimar a sobrevivência, sendo os modelos de Weibull e exponencial os mais usados. Dentre os modelos usados tem-se as seguintes distribuição de probabilidade: Weibull, 23 (51%), exponencial, 20 (44%), Gompertz, 6 (13%), log-logística, 9 (20%), log-normal, 6 (13%) e Gama, com 2 (4%) estudos [17].

1.4.1 O câncer de mama no Brasil e no mundo

Como problema de saúde pública, o câncer de mama se destaca como a neoplasia maligna que mais acomete as mulheres na maior parte do mundo. Em 2018 foram estimados 2,1 milhões de casos novos de câncer, que corresponde a 11,6% dos cânceres estimados e 627 mil óbitos pela doença [18]. No Brasil, as estimativas de incidência de câncer de mama para cada ano do triênio 2020-2022 são em torno de 66 mil casos novos, o que representa 29,7% dos cânceres em mulheres. Em 2017, ocorreram 16.724 mortes de mulheres por câncer de mama no país[19].

Com sua diretriz terapêutica mais recente para o SUS aprovada em 2019 após recomendação da Conitec (Relatório 439 - Diretrizes Diagnósticas e Terapêuticas do Carcinoma de Mama)[20], o câncer de mama é um tema com avaliações frequentes sobre a incorporação ou não de tecnologias pela Conitec e com impacto relevante para a carga das doenças em nosso contexto. Sendo assim, considera-se que o câncer de mama seja um tema de grande interesse das decisões para a aplicação de novas propostas metodológicas relacionadas à ATS

2 OBJETIVO GERAL

Reproduzir um método para extrapolação com curvas de sobrevivência aplicável a dados agregados.

2.1 Objetivos específicos

Levantar as recomendações metodológicas no campo da ATS relacionadas à extrapolação de dados da pesquisa clínica com curvas de sobrevivência.

Aplicar e validar a extração de dados de sobrevivência por meio da digitalização de imagens.

Identificar um método de ajuste de curvas de sobrevivência aplicável a dados agregados.

Aplicar o método em um estudo de caso do câncer de mama metastático.

3 MÉTODOS

O presente trabalho foi dividido em dois momentos. Primeiramente, foram levantadas e consolidadas as principais recomendações disponíveis sobre a extrapolação de dados de sobrevivência para modelos de avaliação econômica. Em um segundo momento, utilizou-se um estudo de caso com o intuito de analisar a aplicabilidade do ajuste de modelos de sobrevivência a partir de dados agregados de ensaios clínicos. A seguir, descrevem-se os passos metodológicos desses dois momentos.

3.1 Revisão das diretrizes disponíveis

3.1.1 Tipo de estudo

No presente estudo, utilizou-se o método de Revisão Rápida, que representa uma estratégia de síntese das evidências relevantes para tomada de decisão de forma ágil[21]. Semelhante às revisões sistemáticas tradicionais, as revisões rápidas são geradas por meio de um método transparente e reproduzível, contudo, simplificando ou abrindo mão de alguns requisitos de uma revisão sistemática em função de obter uma resposta em cenários de tempo e recursos limitados [21]. Na presente revisão, adotou-se um escopo e estratégia de busca restrita. Da mesma forma, as etapas de seleção não foram conduzidas de forma pareada.

3.1.2 Estratégia de busca

A estratégia de busca foi composta por três vertentes: busca em base indexada, busca ativa em agências de ATS e contato eletrônico. Para a busca na base da Medline (via Pubmed), adotou-se a combinação de termos descrita no Quadro 1.

Quadro 1 – Estratégia de busca e resultados selecionados

Base	Estratégia
Medline/Pubmed	("guideline"[Publication Type] OR "guidelines as topic"[MeSH Terms] OR "guideline*" [TIAB] OR "recommendation*" [TIAB]) AND ("technology assessment, biomedical"[MeSH Terms] OR "health technology assessment" [TIAB] OR "Cost-Benefit Analysis" [Mesh] OR "modelling" [TIAB] or "Cost-Effectiveness" [TIAB] OR "Cost-Utility" [TIAB]) AND ("survival" [TIAB] OR "extrapolation*" [TIAB] OR "Survival Analysis" [Mesh])

A segunda fonte foi composta por buscas ativas nos sítios eletrônicos das 10 principais agências de ATS internacionais, conforme escopo delimitado em trabalho anterior de Novaes & Soares (2016), listadas no Quadro 2.

Quadro 2 – Principais Agências de ATS

Agência de ATS	Sigla	País	Sítio
Pharmaceutical Benefits Advisory Committee	PBAC	Austrália	http://www.pbs.gov.au/pbs/home
Swedish Council on Technology Assessment in Health Care	SBU	Suécia	http://www.sbu.se
Canadian Agency for Drugs and Technologies in Health	CADTH	Canadá	http://www.cadth.ca
National Institute for Clinical Excellence	NICE	Reino Unido	www.nice.org.uk
Agency for Healthcare Research and Quality	AHRQ	Estados Unidos	https://www.ahrq.gov/
Institute for Quality and Efficiency in Health Care	IQWiG	Alemanha	http://www.iqwig.de
Centro Nacional de Excelencia Tecnológica en Salud	CENETEC	México	http://www.cenetec.salud.gob.mx
Haute Autorité de Santé	HAS	França	http://www.has-sante.fr
Comissão Nacional de Incorporação de Tecnologias no SUS	CONITEC	Brasil	http://www.conitec.gov.br/
Instituto de Evaluación Tecnológica en Salud	IETS	Colômbia	http://www.iets.org.co

Por fim, foram enviadas mensagens eletrônicas solicitando a colaboração das agências membros da rede internacional de ATS INAHTA (do inglês, International Network of Agencies for Health Technology Assessment) sobre diretrizes metodológicas de extrapolação de dados, conforme comunicação disponível no APÊNDICE A.

3.1.3 Critérios de seleção

Foram considerados diretrizes e demais documentos com recomendações explícitas sobre a extrapolação de dados de sobrevivência na avaliação econômica

em saúde. Estudos de todos os períodos foram aceitos nesta busca, contudo, a seleção foi restrita aos idiomas inglês, espanhol e português.

3.1.4 Extração

Após a inclusão apenas dos estudos que citavam algum método para extrapolação de dados em análise de sobrevivência, adotou-se o formulário padrão apresentado no Apêndice B para a extração de dados.

3.1.5 Síntese

Para a síntese das diretrizes encontradas, buscou-se resumir os dados de acordo com suas características descritivas. Priorizou-se aqui a síntese qualitativa, por não se tratar de uma avaliação de intervenções, mas de escolhas metodológicas.

3.2 Aplicação das diretrizes disponíveis

3.2.1 O estudo de caso no câncer de mama metastático

Como estudo de caso, adotou-se o exemplo da avaliação comparativa de dois tratamentos no câncer de mama metastático (anastrozol e fulvestranto). Para tanto, como fonte primária de dados, fez-se uso dos dados de um ensaio clínico randomizado comparando a efetividade e segurança do tratamento com anastrozol e sua combinação com fulvestranto. A população estudada foi de mulheres na pós-menopausa com câncer de mama metastático hormônio positivo não tratadas anteriormente, ou seja, que não tivessem feito quimioterapia prévia, terapia hormonal ou imunoterapia para doença metastática, conforme Figura 3 [22].

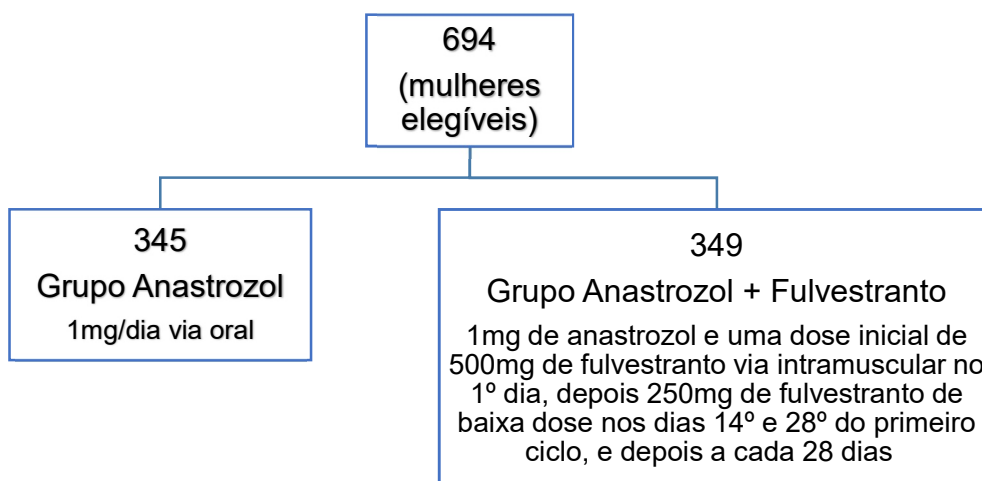


Figura 3 – Divisão dos grupos, para uso dos medicamentos

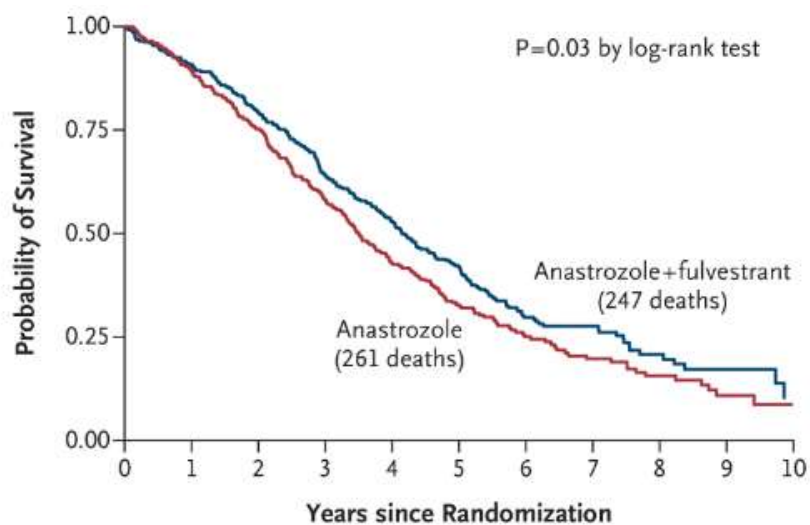
O estudo adotou como desfecho primário a SLP e secundário a SG, fazendo uso do método de K-M para suas estimativas [23]. Para o grupo com tratamento Anastrozol, selecionou-se 345 mulheres elegíveis destas 261 morreram e 84 foram censurados sem morte, para a combinação com o Fulvestranto foram 349 elegíveis, 247 morreram e 102 censurados sem morte, na sobrevida global. Para o grupo com tratamento Anastrozol, 329 tiveram progressão da doença ou morte por qualquer causa e 84 foram censurados sem progressão ou morte, para a combinação com o Fulvestranto, 247 tiveram progressão da doença ou morte por qualquer causa e 102 censurados sem progressão ou morte, na sobrevida livre de progressão.

3.2.2 A digitalização de imagens

Dada a limitação de acesso aos dados individualizados dos estudos, a obtenção dos dados de sobrevivência disponíveis em cada um dos tempos de seguimento foi conduzida com a digitalização dos gráficos das curvas de Kaplan-Meier [24]. Tal abordagem é coerente com as orientações da colaboração Cochrane, que cita a possibilidade de considerar o uso de software para extração dos dados por meio de digitalização de figuras na ausência de dados disponíveis [25]. No presente trabalho utilizamos o software livre WebPlotDigitizer versão 4.2 para realizar a digitalização das imagens, extraíndo os dados das curvas de sobrevivência. O detalhamento da extração e utilização do software encontra-se no APÊNDICE E.

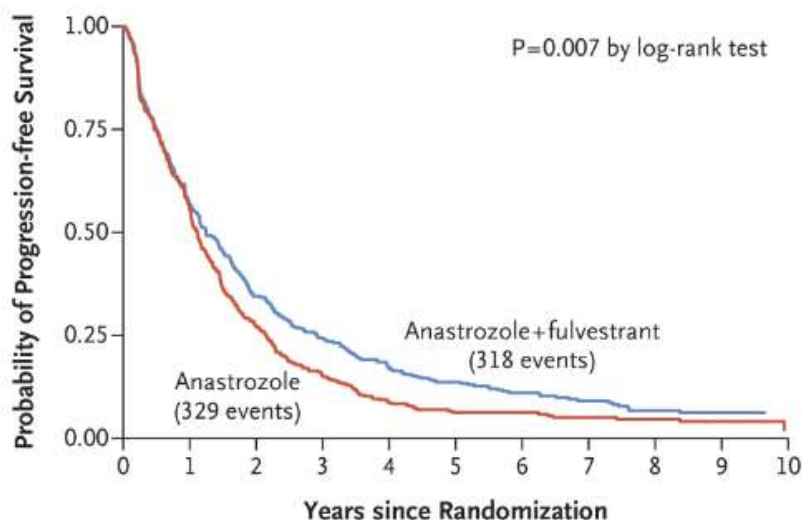
Assim, foram digitalizados os dados das figuras das curvas de K-M de SG (A); e SLP (B) para os dois braços de tratamento publicadas no estudo original, reproduzidas aqui na Figura 4.

(A) - Curva de Kaplan-Meier, Sobrevida Global



No. at Risk	0	1	2	3	4	5	6	7	8	9	10
Anastrozole+fulvestrant	349	315	273	220	177	133	61	38	18	10	3
Anastrozole	345	308	257	196	142	99	45	28	15	7	1

(B) - Curva de Kaplan-Meier, Sobrevida Livre de Progressão



No. at Risk	0	1	2	3	4	5	6	7	8	9	10
Anastrozole+fulvestrant	349	197	119	84	58	45	30	21	13	7	3
Anastrozole	345	192	94	52	29	19	17	13	8	4	1

Figura 4 – Curvas de Kaplan-Meier apresentadas no estudo

Fonte: Mehta, 2019[23]

Adicionalmente, os autores do estudo foram consultados sobre a possibilidade de disponibilização de dados individualizados e desidentificados para fins de comparação das estimativas obtidas com os dados agregados. Na hipótese de obtenção dos dados, foi igualmente prevista a condução de modelos de Kaplan-Meier

e modelos paramétricos a partir de dados individualizados. Após a extração dos dados das imagens, foi plotado gráfico com a interseção das curvas de sobrevivência dos dados individualizados e agregados, para verificar se a retirada das informações da imagem expressa os dados originais, esses pontos devem estar próximos a reta $x=y$, conforme a Figura 5.

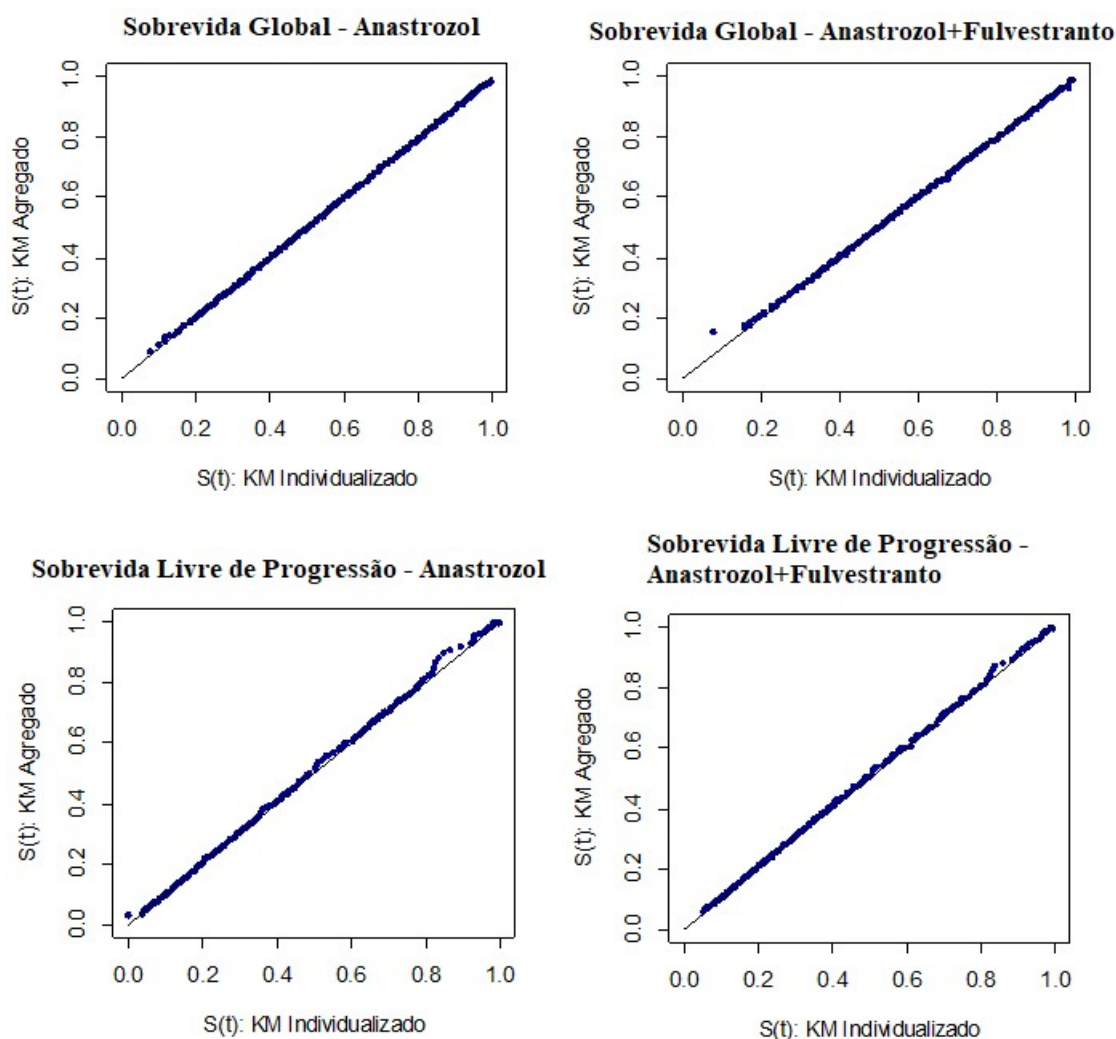


Figura 5 – Curva de sobrevivência de Kaplan-Meier dos dados individualizados versus dados agregados

Na visualização da Figura 5 pode-se observar que os dados extraídos representam bem os dados originais, as curvas apresentam lineariedade em relação a reta $x=y$, ou seja, os dados extraídos das imagens podem ser utilizados para realizar as estimativas. Apesar de um desvio em alguns pontos na sobrevida livre de progressão do Anastrozol, de uma maneira geral os dados obtidos pela digitalização foram bem aderentes aos dados individualizados.

3.2.2.1 Funções de sobrevivência

Com intuito de analisar o ajuste dos modelos aos dados de sobrevivência, foram tomados como base os seis modelos paramétricos citados em recente documento de suporte ao NICE: Exponencial, Weibull, Gompertz, Gama Generalizada, Log-normal e Log-logística [17]

A função de sobrevivência é umas das principais funções probabilísticas. Ela é definida como a probabilidade de uma observação não falhar até um certo tempo t , ou seja, é a probabilidade de uma observação sobreviver ao tempo t , dada uma variável aleatória T não negativa, do tipo contínua, que representa o tempo de falha. De forma probabilística é descrito da seguinte forma e ilustrado na Figura 6:

$$S(t) = P(T > t) \quad (1)$$

Sendo a função de distribuição acumulada o complementar da $S(t)$, pode ser definida como a probabilidade de uma observação sobreviver até o tempo t , como mostra a função abaixo:

$$F(t) = 1 - S(t) \quad (2)$$

A Figura 6 ilustra uma curva teórica da função de sobrevivência, da qual podemos extrair informações sobre a população estudada, como o percentual de pacientes que ainda estão vivos até um tempo de interesse, na curva abaixo poderíamos dizer que no tempo 35, 50% dos pacientes ainda estão vivos.

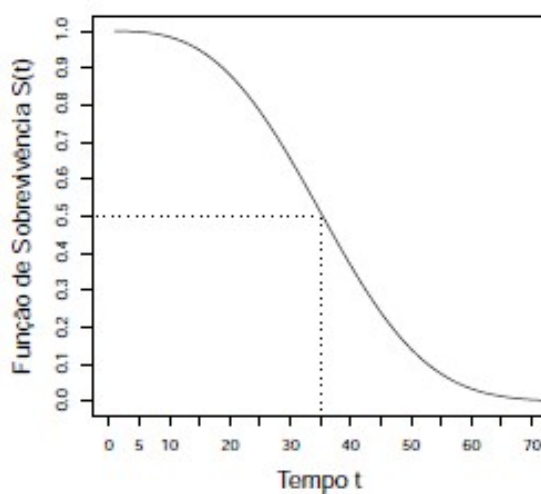


Figura 6 – Curva teórica da função de sobrevivência

Fonte: Carvalho, 2011[14].

3.2.2.2 Função de risco

A função de risco ou taxa de falha $h(t)$, é outra importante função na análise de sobrevivência, pois, ela permite verificar o risco de um indivíduo sofrer a falha em um intervalo de tempo, por exemplo, qual é o risco de um indivíduo com uma determinada doença vir a falecer após sobreviver um ano. Esse risco pode ser constante, crescente ou decrescente, conforme demonstrado teoricamente na Figura 7.

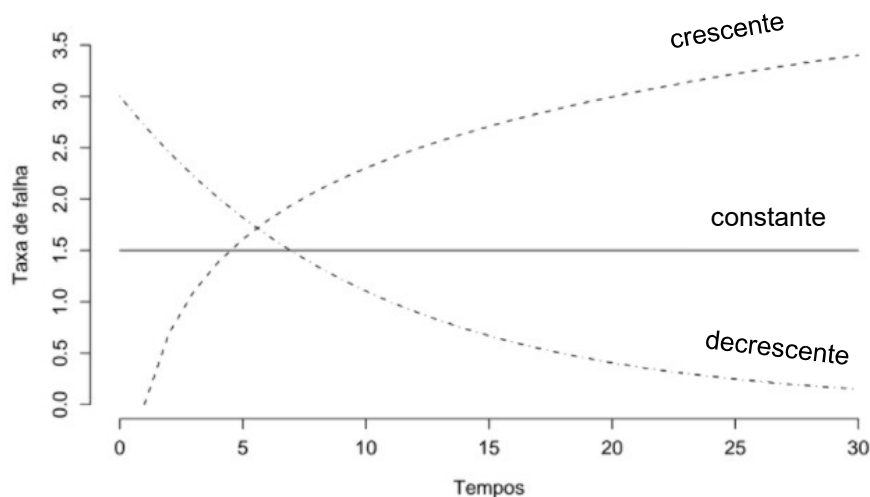


Figura 7 - Curva teórica da função de risco

Fonte: Colosimo, 2006[12]

A função de risco crescente mostra que o indivíduo tem seu risco aumentando no decorrer do tempo. A função constante indica que o risco não se altera no decorrer do tempo. A decrescente diz que o risco diminui no decorrer do tempo.

A função de risco está relacionada com a função de sobrevivência pela expressão:

$$h(t) = \frac{f(t)}{S(t)}$$

3.2.2.2 Distribuição Exponencial

A distribuição exponencial [12] é um dos modelos paramétricos mais simples disponíveis para descrever o tempo de sobrevivência. Possui um único parâmetro e sua função de risco é constante, ou seja, o risco da ocorrência do evento de interesse após o início da observação é o mesmo durante todo o período. Esta condição acaba por embutir uma propriedade de ausência de memória da distribuição.

Para uma variável aleatória de tempo de falha T com distribuição exponencial, sua função densidade de probabilidade (fdp) é dada por:

$$f(t) = \frac{1}{\alpha} \exp\left\{-\left(\frac{t}{\alpha}\right)\right\}, \quad t \geq 0 \quad (3)$$

Onde o único parâmetro $\alpha > 0$ é o tempo médio de vida. E suas funções de sobrevivência $S(t)$ e função de risco $h(t)$ são dadas por:

$$S(t) = \exp\left\{-\left(\frac{t}{\alpha}\right)\right\} \quad (4)$$

$$h(t) = \frac{1}{\alpha}, \quad t \geq 0 \quad (5)$$

3.2.2.3 Distribuição de Weibull

A distribuição de Weibull [19][12] possui uma variedade de formas, todas com uma propriedade básica, sua função de risco $h(t)$ é monótona, ou seja, ela é crescente quando $\gamma > 1$, decrescente quando $\gamma < 1$ ou constante quando $\gamma = 1$, onde tem-se a distribuição exponencial como um caso particular da distribuição Weibull.

Para uma variável aleatória T com distribuição de Weibull, tem-se a função densidade de probabilidade dada por:

$$f(t) = \frac{\gamma}{\alpha^\gamma} t^{\gamma-1} \exp\left\{-\left(\frac{t}{\alpha}\right)^\gamma\right\}, \quad t \geq 0 \quad (6)$$

Na fórmula, o parâmetro de forma é γ , e o de escala é α , ambos positivos. Onde α tem a mesma unidade de medida de t e γ não tem unidade.

Suas funções de sobrevivência $S(t)$ e risco $h(t)$ são dadas por:

$$S(t) = \exp\left\{-\left(\frac{t}{\alpha}\right)^\gamma\right\} \quad (7)$$

e

$$h(t) = \frac{\gamma}{\alpha^\gamma} t^{\gamma-1}, \quad t \geq 0, \quad \alpha \text{ e } \gamma > 0 \quad (8)$$

3.2.2.4 Distribuição Log-normal

A distribuição log-normal [12] também é utilizada para caracterizar tempos de vida em modelos de sobrevivência. A sua função de densidade de probabilidade é dada por:

$$f(t) = \frac{1}{\sqrt{2\pi}t\sigma} \exp\left\{-\frac{1}{2}\left(\frac{\log(t)-\mu}{\sigma}\right)^2\right\}, \quad t \geq 0 \quad (9)$$

Diz-se que uma variável aleatória T tem uma distribuição log-normal, quando o seu logaritmo segue uma distribuição normal, com parâmetros μ e σ , onde μ é a média do logaritmo do tempo de falha e σ é o desvio-padrão.

Suas funções de sobrevivência $S(t)$ e risco $h(t)$, são respectivamente:

$$S(t) = \Phi\left(\frac{-\log(t) + \mu}{\sigma}\right) \quad (10)$$

e

$$h(t) = \frac{f(t)}{S(t)} \quad (11)$$

Em que Φ é a distribuição acumulada de uma normal padrão. A sua função de risco cresce até atingir um valor máximo e depois decrescem.

3.2.2.5 Distribuição de Gompertz

O modelo de Gompertz encontrou aplicação na demografia e nas ciências biológicas, sendo introduzido em 1825 como um modelo para mortalidade humana. A distribuição de Gompertz possui dois parâmetros, o λ de forma e o θ de escala, parecido com a distribuição de Weibull, sua função de risco aumenta ou diminui monotonamente. Sua função de densidade de probabilidade $f(t)$ e sobrevivência são respectivamente, para $t \geq 0$ e λ e $\theta > 0$.

$$f(t) = \lambda e^{\theta t} \exp\left\{\frac{\lambda}{\theta}(1 - e^{\theta t})\right\} \quad (12)$$

e

$$S(t) = \exp\left\{\frac{\lambda}{\theta}(1 - e^{\theta t})\right\} \quad (13)$$

Função de risco:

$$h(t) = \lambda e^{\theta t} \quad (14)$$

3.2.2.6 Distribuição Log-logística

A distribuição log-logística [12] tem função de densidade para uma variável aleatória T que possui distribuição log-logística, com parâmetros α , γ :

$$f(t) = \frac{\gamma}{\alpha^\gamma} t^{\gamma-1} \left(1 + \left(\frac{t}{\alpha}\right)^\gamma\right)^{-2}, t > 0 \quad (15)$$

Sendo $\alpha > 0$ o parâmetro de escala e $\gamma > 0$ o de forma. A distribuição é assim denominada porque a variável $\log T$ possui uma distribuição logística, uma distribuição simétrica cuja função densidade de probabilidade é muito semelhante à da distribuição normal. E as funções de sobrevivência e risco dada por, respectivamente:

$$S(t) = \frac{1}{1 + \left(\frac{t}{\alpha}\right)^\gamma} \quad (16)$$

e

$$h(t) = \frac{\gamma \left(\frac{t}{\alpha}\right)^{\gamma-1}}{\alpha \left[1 + \left(\frac{t}{\alpha}\right)^\gamma\right]} \quad (17)$$

Os modelos log-logísticos muitas vezes resultam em caudas longas na função sobrevivência, e isso deve ser levado em consideração, quando forem usados.

3.2.2.7 Distribuição Gama Generalizada

A distribuição gama generalizada é uma distribuição bastante compreensiva, incluindo como casos particulares as distribuições: exponencial, Weibull, gama e log-normal. Sendo a sua função de densidade $f(t)$:

$$f(t) = \frac{\gamma}{\Gamma(\kappa) \alpha^{\gamma\kappa}} t^{\gamma\kappa-1} \exp\left\{-\left(\frac{t}{\alpha}\right)^\gamma\right\}, t > 0 \quad (18)$$

$$h(t) = \frac{f(t)}{S(t)} \quad (19)$$

Em que $\Gamma(\kappa) = \int_0^{\infty} x^{\kappa-1} \exp\{-x\} dx$ é definida como a função gama. A distribuição possui três parâmetros γ de escala, κ e α de forma, todos positivos.

A partir da função de densidade da distribuição gama generalizada temos como casos particulares:

- 1) para $\gamma = \kappa = 1$ tem-se $T \sim \text{Exp}(\alpha)$,
- 2) para $\kappa = 1$ tem-se $T \sim \text{Weibull}(\gamma, \alpha)$,
- 3) para $\gamma = 1$ tem-se $T \sim \text{Gama}(\kappa, \alpha)$,
- 4) para $\kappa \rightarrow \infty$ tem-se $T \sim \text{log-normal}$.

3.3.1 Ajuste dos modelos

Para a avaliação dos modelos será feita a inspeção visual e comparação entre as estatísticas de critérios de informação: *Akaike's Information Criterion* (AIC) e *Bayesian Information Criterion* (BIC) [12].

No presente trabalho, a inspeção visual adotou dois métodos gráficos. O primeiro faz a comparação das funções de sobrevivência dos modelos em estudo com o estimador de Kaplan-Meier versus o tempo, verificando as distribuições paramétricas que mais se ajustam, isto é, as que mais se aproximam da curva estimada de Kaplan-Meier. Uma outra variação deste método é analisar os pontos das funções de sobrevivência estimada de Kaplan-Meier (no eixo x) versus as funções de sobrevivência das distribuições paramétricas (no eixo y), verificando se os pontos plotados estão correlacionados (próximos a reta $x = y$), que indicará uma adequação da distribuição aos dados.

Os critérios AIC e BIC, são utilizadas para analisar de forma comparativa o ajuste dos modelos, sendo possível avaliá-los levando em consideração não apenas a minimização dos erros como também o número de parâmetros dos modelos. Após estimadas as estatísticas, deve-se escolher os modelos com os menores valores de AIC e BIC[12].

Os passos citados foram implementados com a linguagem de programação livre R por meio da aplicação RStudio versão 1.3.1093. Tendo acesso aos dados individualizados (do inglês, *Individual Patient Data* – IPD) as rotinas foram implementadas com o auxílio dos pacotes estatísticos: Surv, Survfit, Flexsurvreg. Para

os dados agregados, que são dados de sobrevivência, obtidas da curva de Kaplan-Meier da imagem, que não estão disponíveis a nível de indivíduos (*Individual Patient Data* – IPD), buscou-se a aplicação de modelos de otimização com o pacote de ajuste não linear (do inglês, Nonlinear Least Squares – NLS). Os dados agregados obtidos da digitalização das imagens, estão expostos no APÊNDICE E.

4 RESULTADOS

4.1 Revisão Rápida das diretrizes

Após a condução da busca de acordo com os métodos descritos, foram selecionados os estudos identificados de acordo com fluxograma de seleção apresentado na Figura 8.

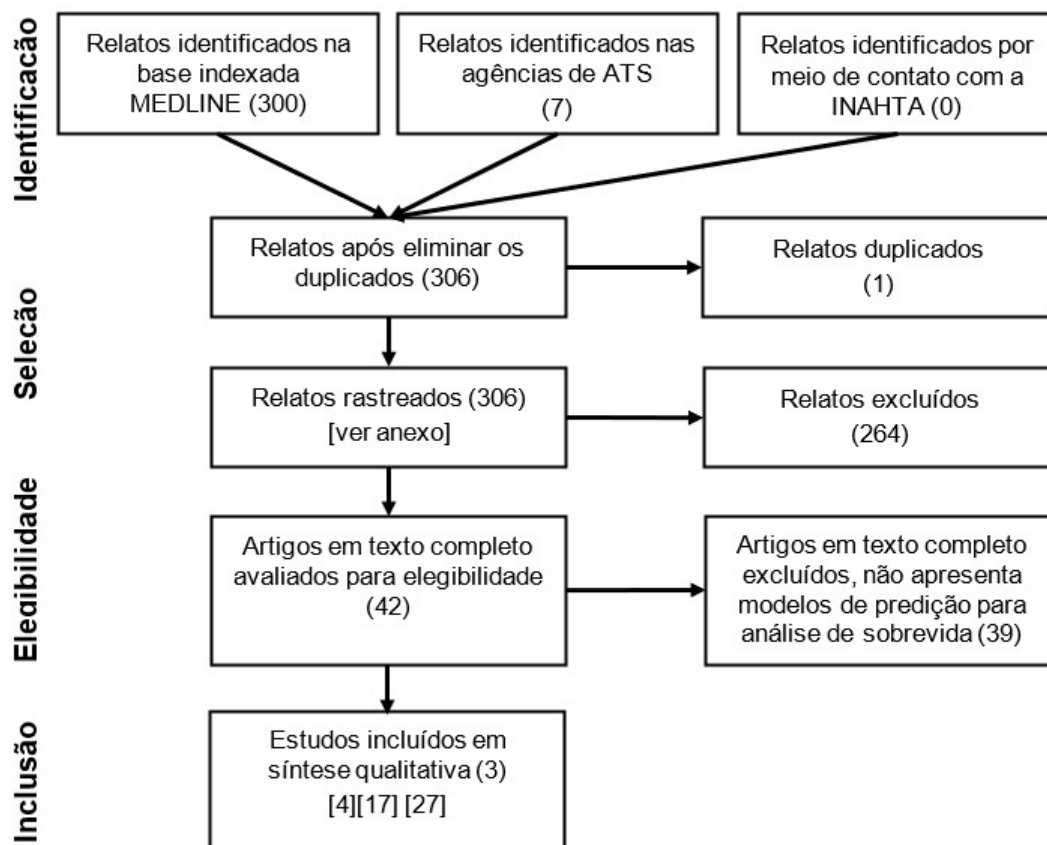


Figura 8 – Fluxograma de seleção das publicações

Após a análise do texto completo dos estudos incluídos, pode-se identificar que as principais distribuições recomendadas foram: Exponencial, Weibull, Log-normal, Log-logística, Gama generalizada e Gompertz conforme resumo no Quadro 3. As orientações técnicas disponíveis só tratavam de dados individualizados. A seguir serão resumidas as diretrizes.

Quadro 3 – Síntese dos resultados obtidos

Estudo	País do Estudo	Ano da publicação	Função recomendada	Crterios de avaliação do ajuste
Pharmaceutical Benefits Advisory Committee (PBAC)	Australia	2016	exponencial, Weibull, log-normal, log-logística, gama, Gompertz	Inspeção visual, critério de informação de Akaike (AIC), Critério de informação bayesiano (BIC)
National Institute for Clinical Excellence (NICE)	Reino Unido	2011	exponencial, Weibull, Gompertz, log-logística, log-normal e gama generalizada	Inspeção visual, Log-cumulative hazard plot, critério de informação de Akaike (AIC), Critério de informação bayesiano (BIC),
Comissão Nacional de Incorporação de Tecnologias no SUS (CONITEC)	Brasil	2014	exponencial, Weibull, Gompertz, log-normal.	Inspeção visual, critério de informação bayesiano e de Akaike (BIC e AIC)

4.1.1 PBAC

O documento de diretrizes do *Pharmaceutical Benefits Advisory Committee* (PBAC) intitulado *Guidelines for preparing a submission to the Pharmaceutical Benefits Advisory Committee* e publicado em 2016 tem o objetivo de direcionar a preparação de uma submissão ao PBAC para avaliação sobre a incorporação de um medicamento no sistema de saúde australiano. Nele, encontram-se, entre outros pontos, quais informações gerais são necessárias, qual a forma mais apropriada de evidência clínica e avaliação econômica.

A diretriz está dividida em cinco seções: a primeira seção fala do contexto e apresentação do medicamento, onde se informa como devem ser fornecidos detalhes essenciais do medicamento e sua proposta de uso. A seção seguinte indica como apresentar a melhor evidência clínica disponível juntamente com a eficácia e segurança do medicamento proposto e a indicação ao paciente. A terceira seção mostra quais os passos a seguir para realizar uma avaliação econômica que possibilite substituir o medicamento proposto pelo comparador. Nesse ponto, ressalta-se que deve ser feita uma descrição completa e transparente da avaliação econômica, incluindo análises de sensibilidade para dimensionar a incerteza em torno dos resultados.

Nesta última seção, onde se discute o uso da avaliação econômica, é indicada a possibilidade de extrapolação de dados, usando dados de pacientes individuais, se estes estiverem disponíveis. Ao trazer tal possibilidade, é exemplificado o uso de um

conjunto padronizado de distribuições paramétricas: exponencial, Weibull, log-normal, log-logística, gama e Gompertz. Além de citar o conjunto de distribuições, o texto ressalta a necessidade de ser avaliado o ajuste das curvas extrapoladas com base em:

- ✓ Inspeção visual;
- ✓ Critério de informação de Akaike (AIC); e
- ✓ Critério de informação bayesiano (BIC).

A penúltima seção aborda a análise de impacto orçamentário, estimando a provável adoção do medicamento proposto na prática clínica e o impacto no custo do orçamento do governo australiano. Por fim, o documento traz questões adicionais que sejam relevantes para a submissão ao PBAC.

4.1.2 NICE

Foi incluído na síntese um dos documentos da unidade de suporte técnico (Decision Support Unit - DSU) da agência inglesa de ATS NICE que tem por título: *Survival Analysis For Economic Evaluations Alongside Clinical Trials - Extrapolation With Patient-Level Data*, autor Nicholas Latimer, publicado em 2011, no Reino Unido. Tal document, produzido pela Universidade de Sheffield, no reino Unido, trata de uma revisão de métodos e proposta para uso da análise de sobrevivência de forma sistemática em modelos econômicos.

O documento, dividido em sete capítulos, traz uma introdução e relevância da análise de sobrevivência nas avaliações de Tecnologias submetidas ao NICE. Na sequência, são apresentados os métodos de modelagem de análise de sobrevivência, englobando suas diferenças de outros tipos de dados contínuos, como o fato de que os desfechos de interesse muitas vezes não são observados em todos os indivíduos (censura).

São abordados os modelos paramétricos e sua capacidade de extrapolação dos dados. Neste ponto, é também apresentada uma gama de distribuições que podem ser usadas, cada uma com suas características: exponencial, Weibull, Gompertz, log-logística, log-normal e gama generalizada.

Ao se discutir a avaliação da adequação dos modelos no que diz respeito ao ajuste aos dados observados e se a parcela que foi extrapolada é clinicamente plausível, são também apresentados alguns critérios:

- ✓ Inspeção visual;
- ✓ Gráfico do log do *hazard* acumulado;
- ✓ Critério de informação AIC e BIC

Adicionalmente, o documento adverte que tais mecanismos possuem limitações, pois, essas avaliações levam em consideração apenas o ajuste dos dados extrapolados com os dados observados, não informando o quão adequado é um modelo paramétrico além dos dados usados.

Ao conduzir uma revisão dos métodos de análise de sobrevivência que foram utilizados nas avaliações de tecnologias, foram revistos 45 documentos, sobre câncer, desde dezembro de 2009 apresentadas ao NICE. O intuito era de verificar os métodos de distribuições de sobrevivência utilizados. Dos quarenta e cinco documentos 32 (71%) utilizaram métodos paramétricos, com as distribuições de Weibull 23 (51%) exponencial 20 (44%) Gompertz 6 (13%) log-logística 9 (20%) log-normal 6 (13%) Gama 2 (4%) *Piecewise modelling* 1 (2%). Para validação dos modelos usados, 22 dos 32 estudos forneceram algum motivo para a escolha da distribuição. Sendo a inspeção visual umas das únicas justificativas para a escolha do modelo em várias avaliações de tecnologias, sendo raro o teste da validade interna dos modelos paramétricos aos dados observados.

Com base nesta revisão e análise, o documento traz uma orientação metodológica e de processo, para que a seleção dos modelos seja feita de forma transparente e sistemático conforme a Figura 9. O último capítulo faz menção as referências utilizadas para elaboração do documento de suporte técnico.

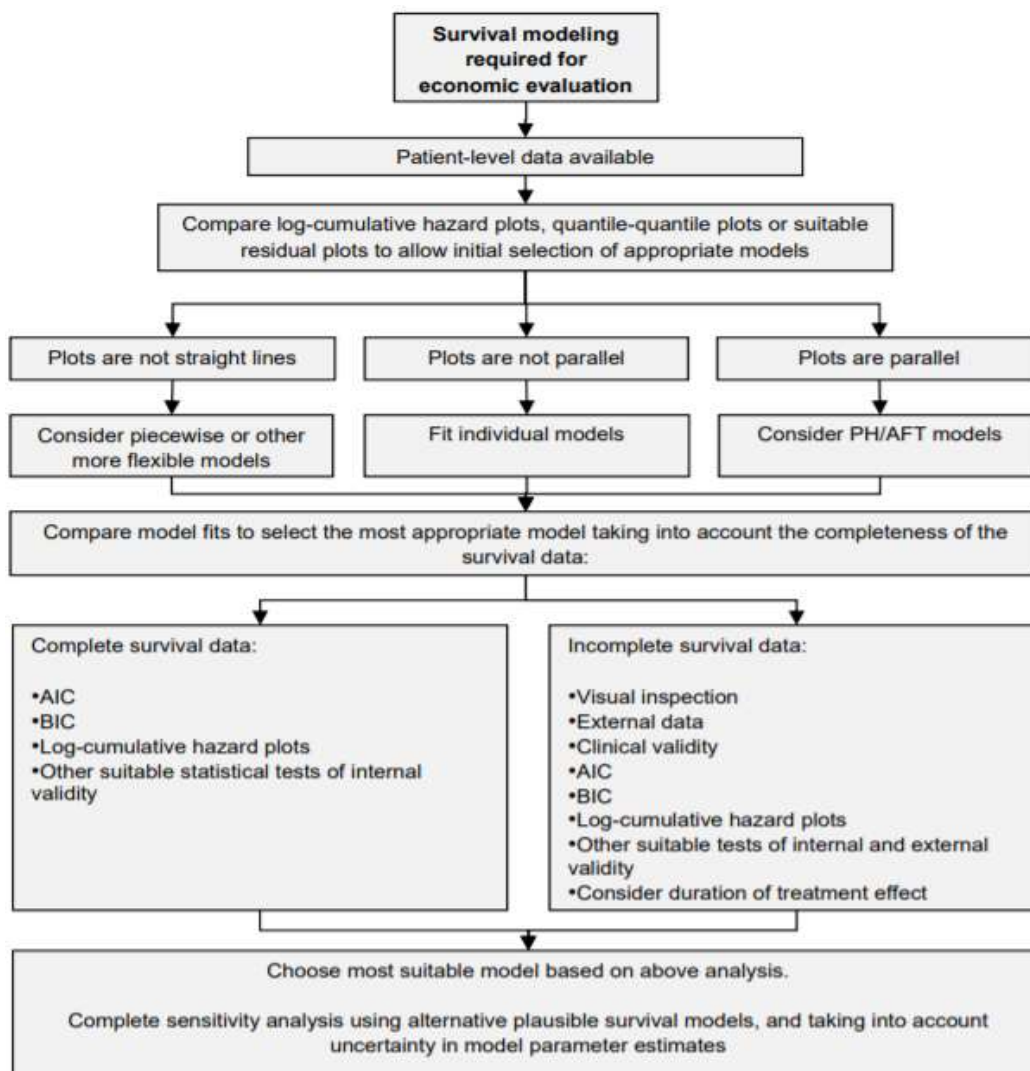


Figura 9 – Processo para escolha do modelo de sobrevivência

Fonte: NICE, 2011[17]

O fluxo da Figura 9 exemplifica o processo com um passo a passo para a escolha do modelo de análise de sobrevivência, sendo as seguintes etapas evidenciadas:

- Uso de dados individualizados;
- Verificar a adequação dos modelos usando uma inspeção visual utilizando o gráfico do log dos riscos acumulados, se a curva produzida for aproximadamente reta, então o modelo está adequado, caso contrário, usar outros modelos mais flexíveis. Se a curva para os braços de tratamento não forem paralelas, então deve-se ajustar um modelo para cada braço de

tratamento e se as curvas forem paralelas pode-se considerar modelos de riscos proporcionais.

- Compara-se os modelos que foram visualmente ajustados, aplicando o teste de critério de informação AIC/BIC ou outro teste adequado. Se os dados não sofreram muitas perdas a escolha do modelo pode se basear nestes testes. Caso contrário, se houver uma censura significativa deve-se levar em consideração dados externos, plausibilidade clínica e avaliação especializada para verificar a validade interna e externa do modelo.

4.1.3 CONITEC

O documento incluído para a síntese dos resultados intitulado Diretrizes metodológicas: Diretriz de Avaliação Econômica, publicado pelo Ministério da Saúde em 2014. É uma diretriz para a condução de modelos de avaliação econômica em saúde, sobretudo, que serão submetidos à Conitec para incorporação de uma tecnologia no SUS.

O documento situa sobre o histórico da criação dos setores que são responsáveis por avaliações de tecnologias em saúde no país. E a necessidade de se realizar os estudos para incorporação de uma nova tecnologia, juntamente com suas avaliações econômicas.

Dentro das recomendações gerais para realização de uma avaliação econômica, são abordados todos os aspectos que um estudo de avaliação econômica deve ter, como: caracterização do problema a ser estudado, a população que será considerada, o desenho do estudo, o tipo de análise (avaliação econômica parcial ou completa), a importância do horizonte temporal entre outros. Ao orientar sobre a projeção de sobrevivência, a diretriz ressalta a limitação de que a maioria dos estudos possui um seguimento curto, não permitindo estimativas de longo prazo, apenas baseando-se nos dados disponíveis. Neste momento, o documento destaca a possibilidade do uso de modelos de predição com curvas de sobrevivência. São também sugeridas como padrão algumas distribuições de probabilidade como: exponencial, Weibull, Gompertz e log-normal.

Semelhante às demais recomendações, o documento do Ministério da Saúde brasileiro orienta que a escolha da distribuição deve se dar por:

- ✓ Inspeção visual;
- ✓ Critério de informação de Akaike (AIC); e
- ✓ Critério de informação bayesiano (BIC).

Estes critérios indicarão quais dentre as funções melhor se ajusta aos dados.

Finalizando o documento direciona sobre o formato de apresentação dos estudos, ele informa como os itens devem ser incluídos ao se relatar os resultados de avaliações econômicas de intervenções para a saúde.

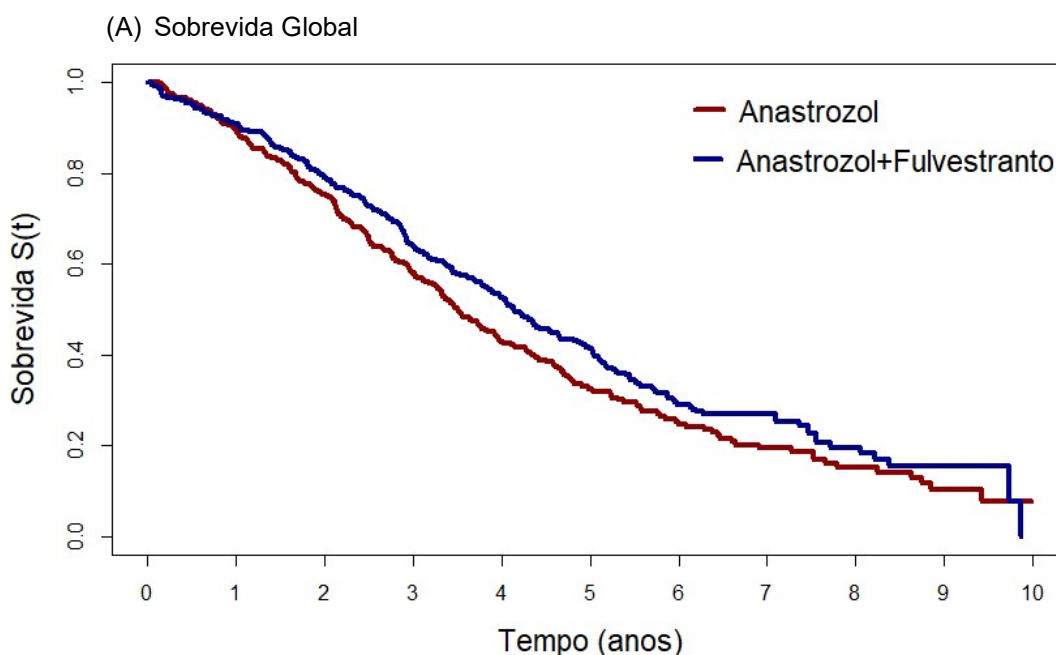
4.2 Ajuste de curvas aos dados no câncer de mama

4.2.1 Dados de sobrevivência observada

Estimativas com dados individualizados

Após o contato com o autor Dr. William Barlow (Universidade de Washington - EUA), foram disponibilizados os dados individualizados do estudo "*Overall Survival with Fulvestrant plus Anastrozole in Metastatic Breast Cancer*".

Ao conduzir a análise de sobrevivência com um estimador de Kaplan-Meier com os dados individualizados disponibilizados, obteve-se as curvas apresentadas no Gráfico 1, referentes à SG (A) e SLP (B).



(B) Sobrevida Livre de progressão

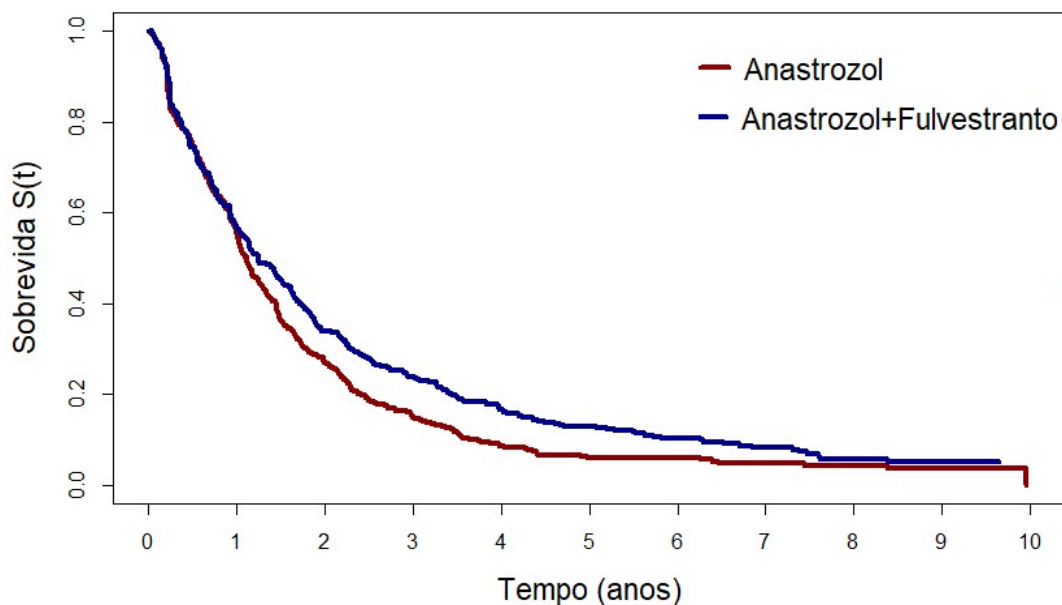


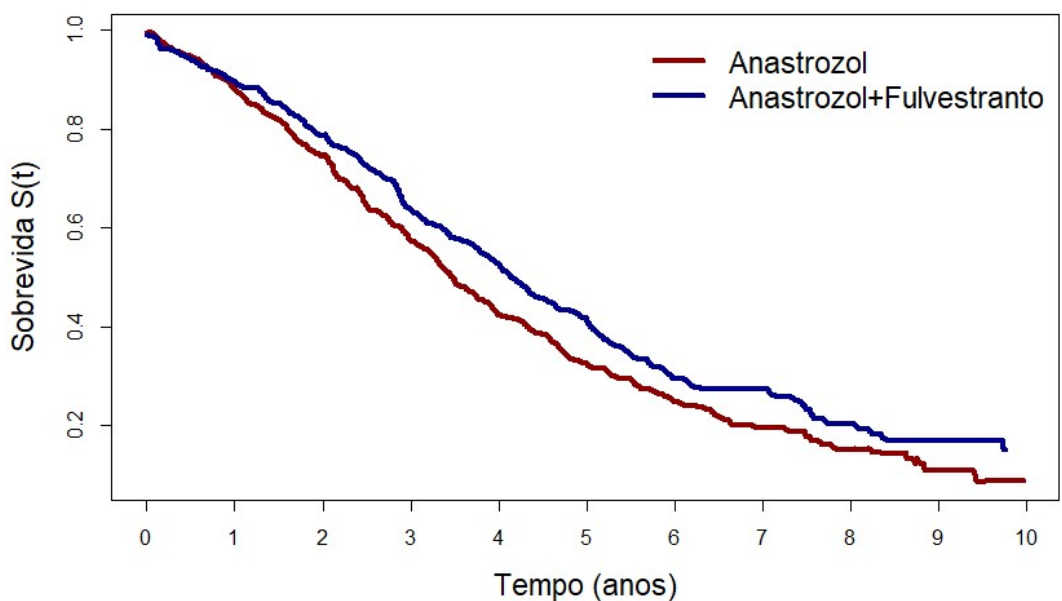
Gráfico 1 – Curvas de Kaplan-Meier obtidas dos dados individualizados para os dois braços de tratamento

Fonte: Elaboração própria a partir dos dados encaminhados pelo Dr. William Barlow (Universidade de Washington -EUA).

Estimativas com dados agregados

Após a digitalização das imagens disponibilizadas no artigo científico seguem as tábuas de vida com os dados agregados na forma de percentuais de sobrevivência ao longo do tempo, conforme Gráfico 2.

(A) Sobrevida global



(B) Sobrevida livre de progressão

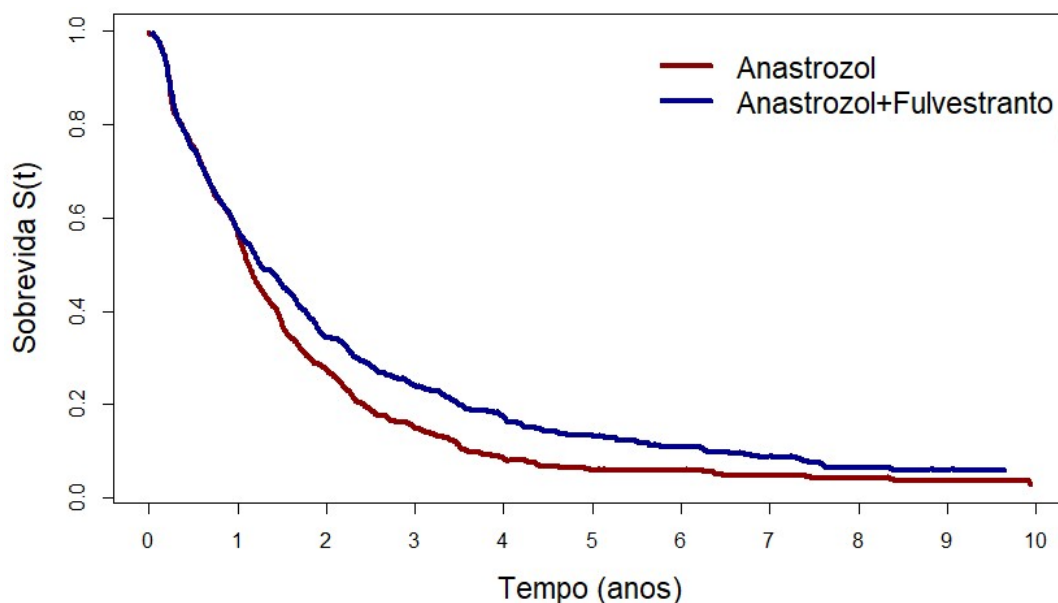


Gráfico 2 – Curvas obtidas das figuras digitalizadas para os dois braços de tratamento

Fonte: Elaboração a partir dos dados disponibilizados, 2019 [23].

4.2.2 Dados de sobrevivência ajustados

Estimativas com dados individualizados

Seguindo o roteiro proposto pelas diretrizes identificadas, foram propostos modelos paramétricos (exponencial, Weibull, log-logística, log-normal, gama generalizada e Gompertz) e sua adequação foi avaliada com a inspeção visual e as estatísticas de ajuste dos critérios (AIC e BIC)

Inspeção visual

Os braços de tratamentos foram avaliados separadamente conforme Gráfico 3 e Gráfico 4 para o ajuste das curvas para o Anastrozol e Anastrozol+Fulvestranto com os dados individualizados para sobrevida global(A) e sobrevida livre de progressão(B).

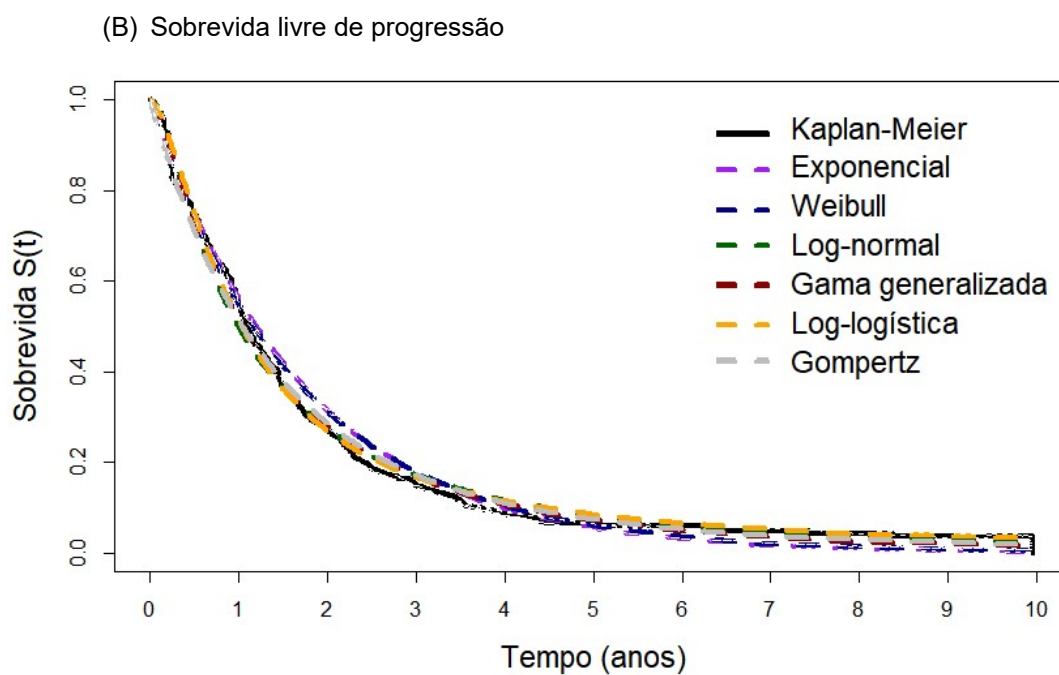
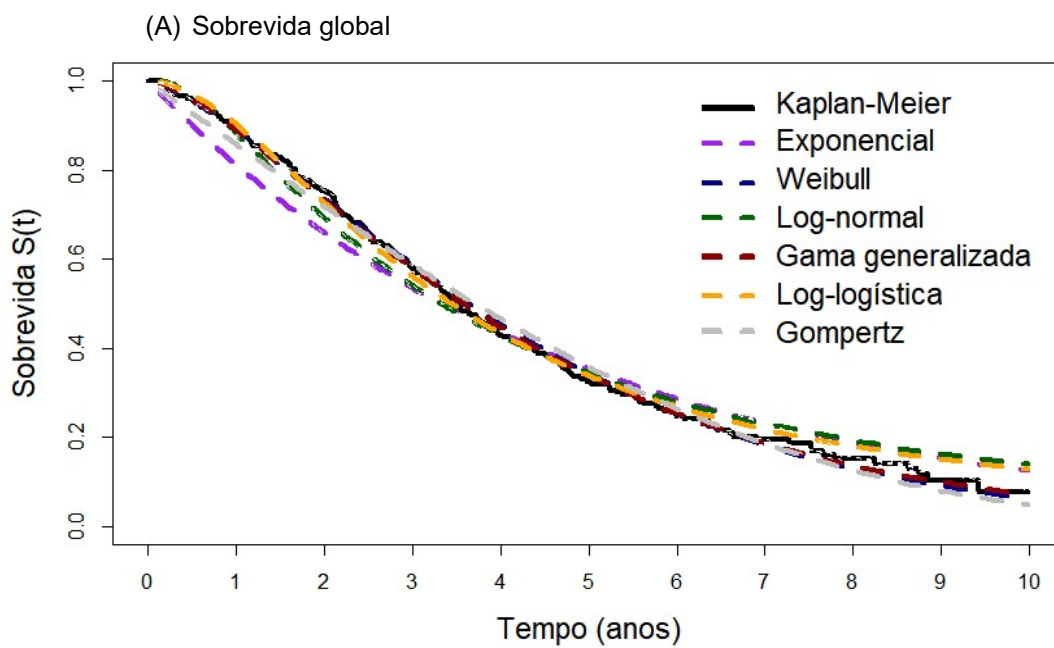


Gráfico 3 – Curvas de Kaplan-Meier obtidas dos dados individualizados para o Anastrozol vs distribuições estimadas.

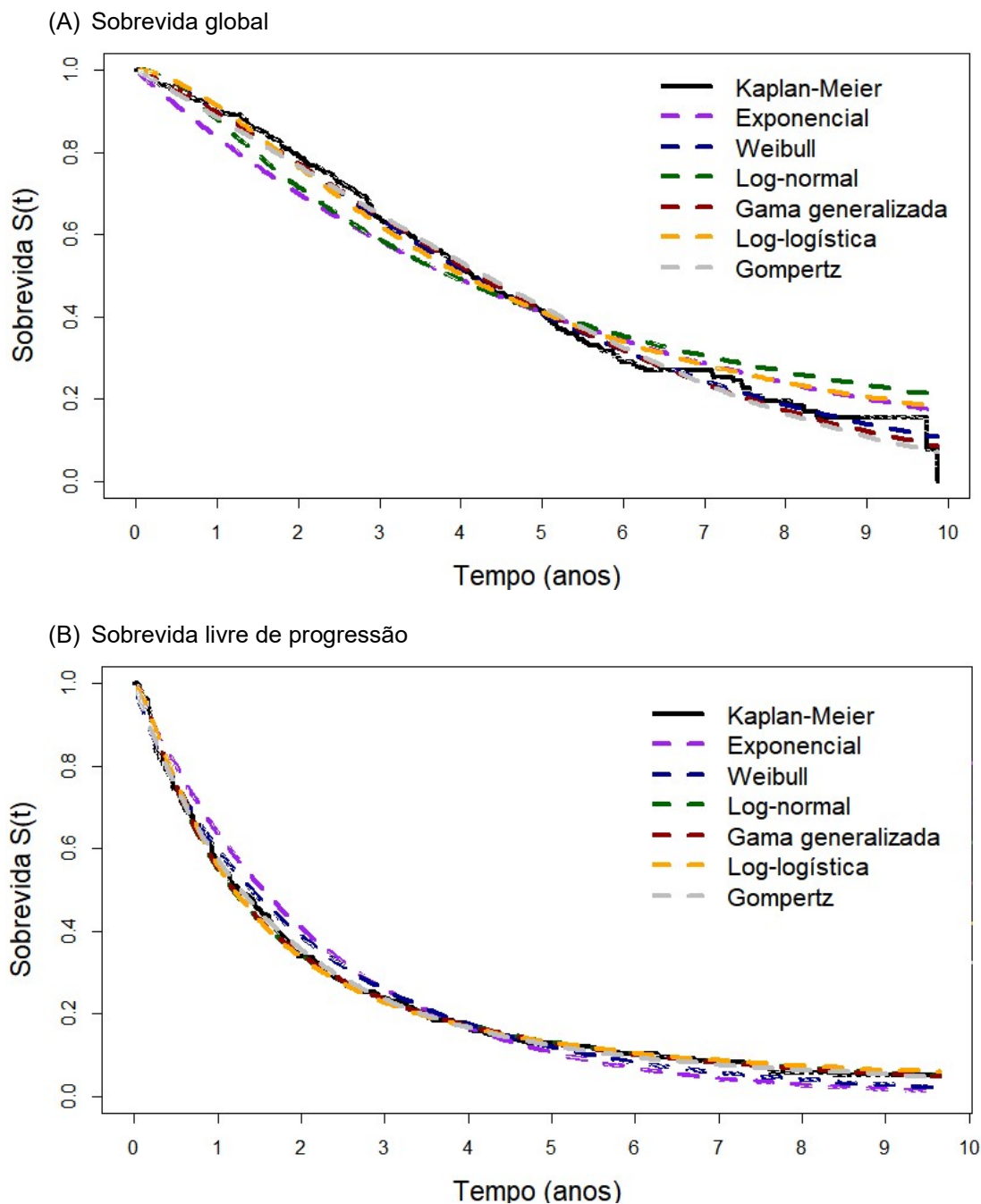


Gráfico 4 – Curvas de Kaplan-Meier obtidas dos dados individualizados para o Anastrozol+Fulvestranto vs distribuições estimadas.

Para os dois braços de tratamentos as curvas estimadas que melhor se ajustaram visualmente aos dados foram as distribuições Weibull e gama generalizada com adição da Gompertz para o tratamento combinado com o fulvestranto, comparado com a curva de Kaplan-Meier para sobrevida global(A). Para a sobrevida livre de progressão(B) nos dois braços de tratamento quase todas as curvas ficaram bem próximas da curva de Kaplan-Meier, as duas que ficaram menos próximas foram

exponencial e Weibull. Os parâmetros calculados das distribuições estão disponibilizados no APÊNDICE F.

Foram plotadas as curvas de sobrevivência estimadas dos modelos propostos versus a curva de sobrevivência obtida pelo método de Kaplan-Meier para analisarmos as curvas estimas por outro método de inspeção visual, temos a Figura 10 e Figura 11.

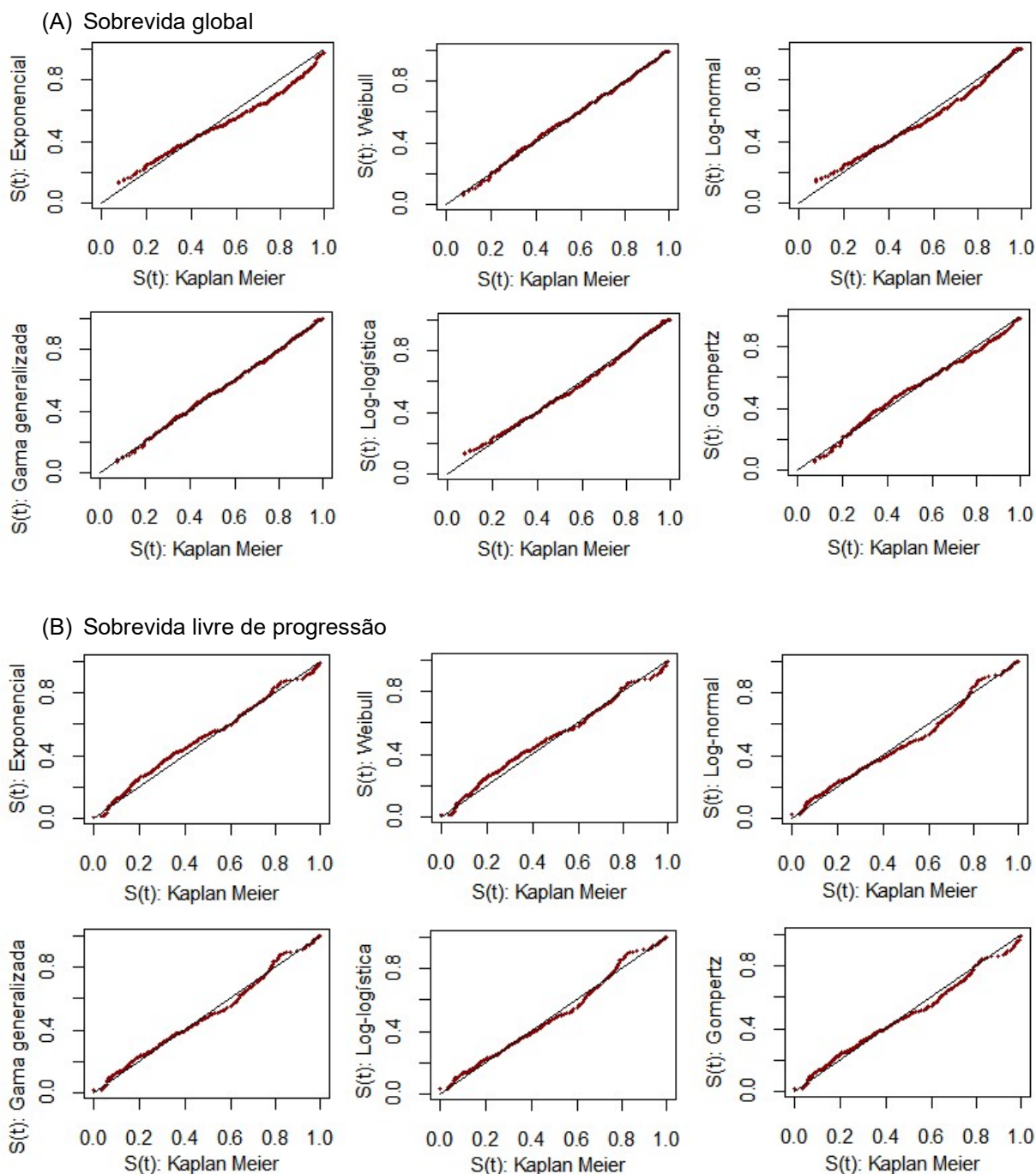
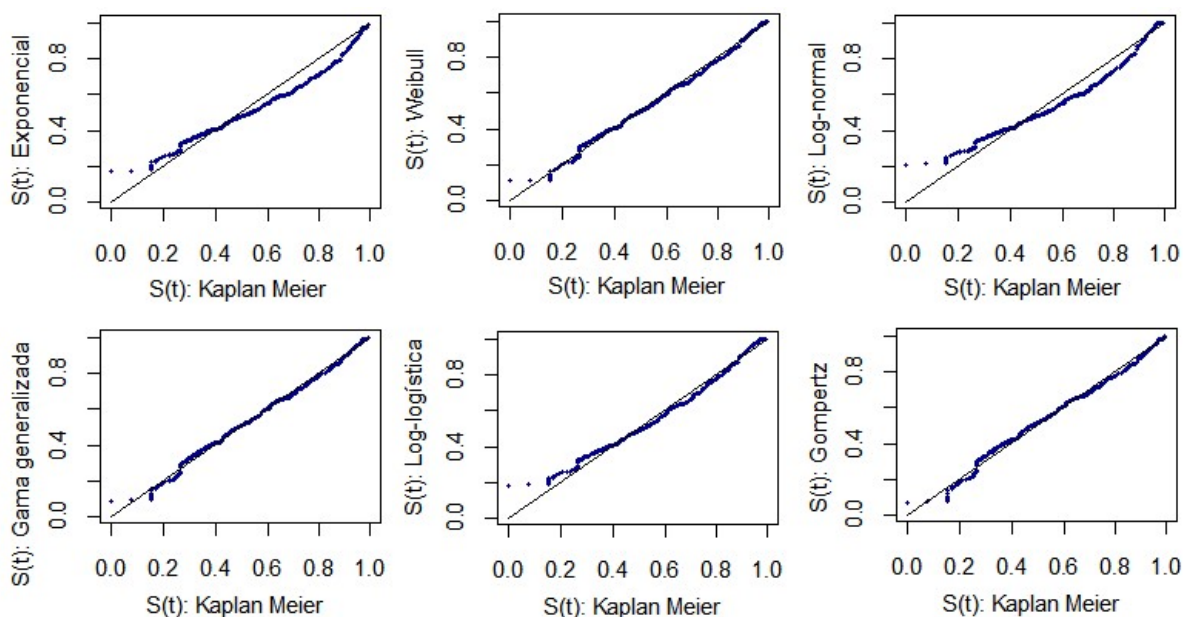


Figura 10 – Curva de sobrevivência de Kaplan-Meier versus curvas de sobrevivência estimadas propostas para os dados individualizados – tratamento Anastrozol, sobrevida global e livre de progressão

(A) Sobrevida global.



(B) Sobrevida livre de progressão

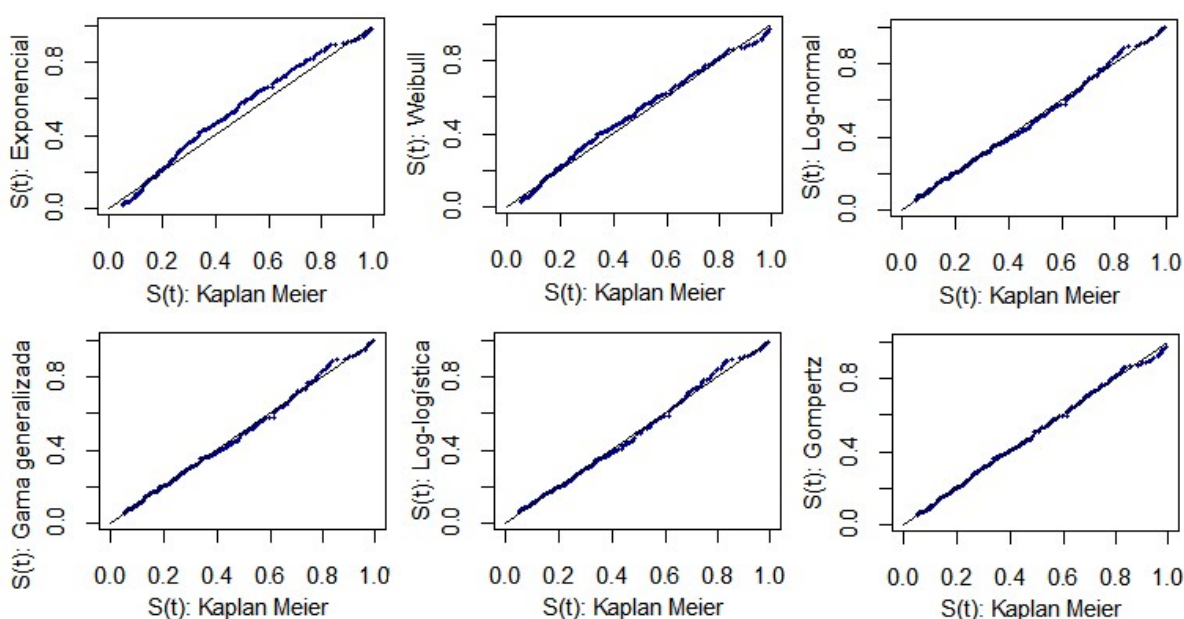


Figura 11 – Curva de sobrevivência de Kaplan-Meier versus curvas de sobrevivência estimadas propostas para os dados individualizados – tratamento Anastrozol+Fulvestranto, Sobrevida global e livre de progressão

De forma consistente com a primeira inspeção visual baseada nas curvas de sobrevivência, neste outro método de inspeção verifica-se também que em ambos os braços de tratamento as curvas estimadas Weibull e gama generalizada são as que mais se ajustam a reta $y=x$, a distribuição de Gompertz também parece aderente no tratamento combinado com fulvestranto, mostrando que estes modelos podem ser adequados aos dados do estudo, conforme Figura 10 e Figura 11 para a sobrevida

global(A). Para a sobrevida livre de progressão(B), podemos destacar que as que ficaram menos próximas a reta $x=y$ foram Weibull e exponencial. Verifica-se que a curva gama generalizada mostrou-se uma candidata tanto para ambos os tratamentos quanto para as sobrevidas.

Estatísticas de Ajuste

Ao considerar as estatísticas de ajuste AIC e BIC, verificou-se que as distribuições Weibull em primeiro lugar e gama generalizada em segundo possuem o menor valor para os dois critérios de avaliação para o tratamento com o Anastrozol, para o tratamento com a combinação do Fulvestranto tem-se que as distribuições com menores valores das estatísticas de ajuste são Gama generalizada (AIC), Weibull (BIC), para a sobrevida global. Ao analisar a sobrevida livre de progressão verifica-se que para o tratamento Anastrozol+Fulvestranto as curvas log-normal e gama generalizada se ajustaram aos dados nos dois critérios de informação (AIC e BIC), já para o tratamento com Anastrozol foi a curva log-logística e gama generalizada para o critério AIC e log-normal para o critério BIC, conforme verificado na Tabela 1.

Tabela 1: Estatísticas do critério de informação AIC e BIC, sobrevida global e livre de progressão, para os tratamentos Anastrozol e Anastrozol+Fulvestranto, para os dados individualizados

Distribuição	Sobrevida Global				Sobrevida livre de progressão			
	Anastrozol		Anastrozol+Fulvestranto		Anastrozol		Anastrozol+Fulvestranto	
	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
Exponencial	1342.930	1346.771	1346.975	1350.822	1018.7919	1022.633	1147.795	1151.641
Weibull	1313.428	1321.109	1323.797	1331.490	1019.9017	1027.583	1141.275	1148.968
Log-normal	1334.343	1342.024	1376.212	1383.904	998.9785	1006.660	1112.601	1120.294
Gama Generalizada	1314.779	1326.301	1323.454	1334.994	996.9148	1008.437	1114.537	1126.077
Log-logística	1319.823	1327.504	1341.979	1349.672	996.1412	1003.822	1120.047	1127.740
Gompertz	1325.198	1332.879	1323.652	1331.345	1005.5318	1013.213	1126.659	1134.352

Nota: AIC: Akaike's Information Criterion e BIC: Bayesian Information Criterion

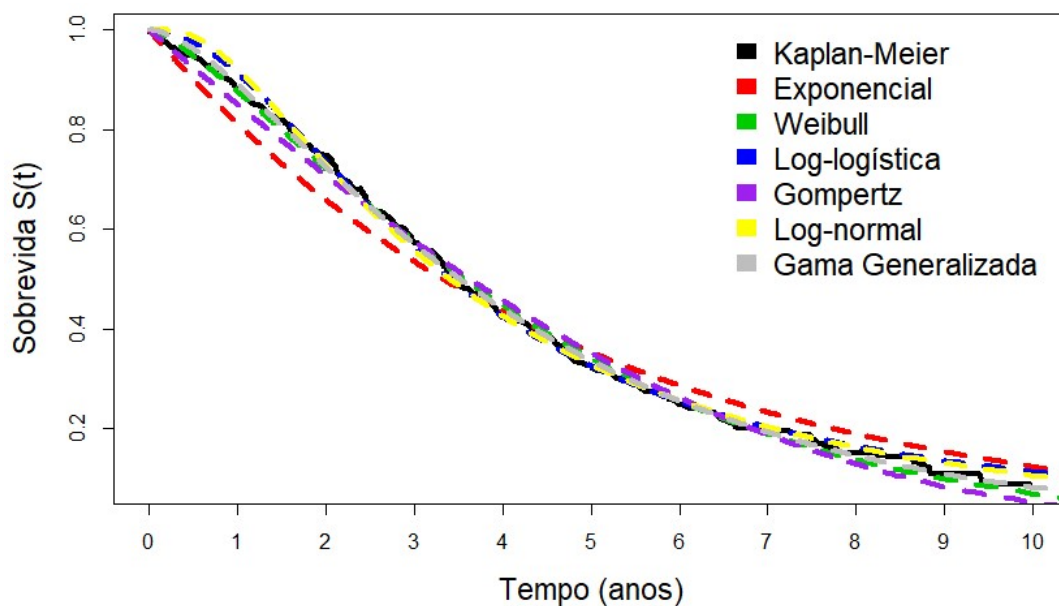
Estimativas com dados agregados

A seguir são apresentados os resultados da inspeção visual e estatísticas de ajuste baseados nos dados agregados.

Inspeção visual

Os dois braços de tratamentos foram avaliados separadamente conforme Gráfico 5 e Gráfico 6 usando os dados agregados. Neste método de inspeção analisamos as curvas estimadas de sobrevivência com a curva de Kaplan-Meier ao longo do tempo.

(A) Sobrevida global



(B) Sobrevida livre de progressão

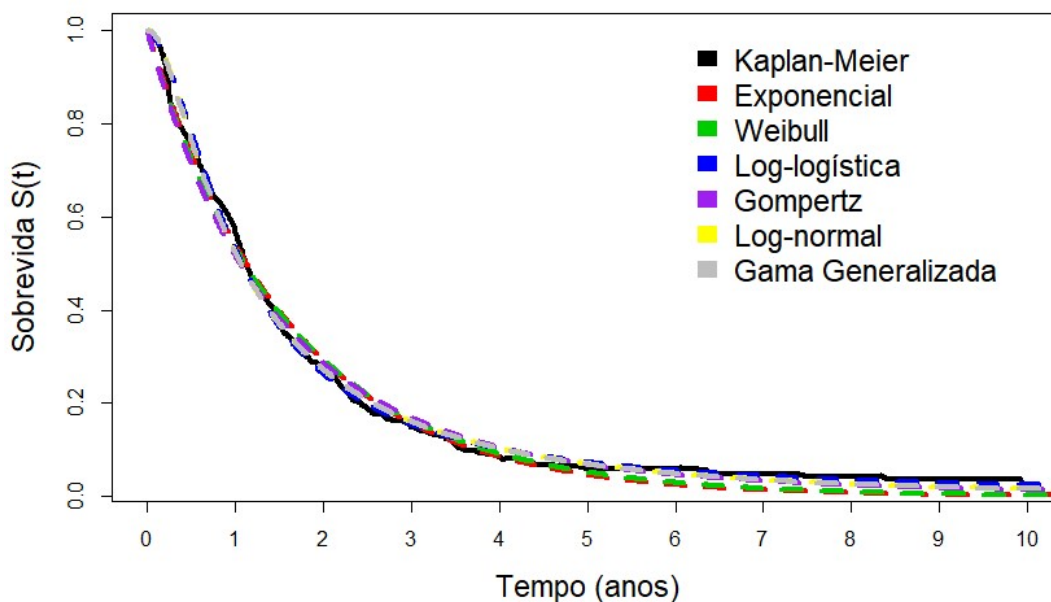


Gráfico 5 – Curvas da Kaplan-Meier obtidas dos dados agregados para o Anastrozol vs distribuições estimadas – sobrevida global e livre de progressão

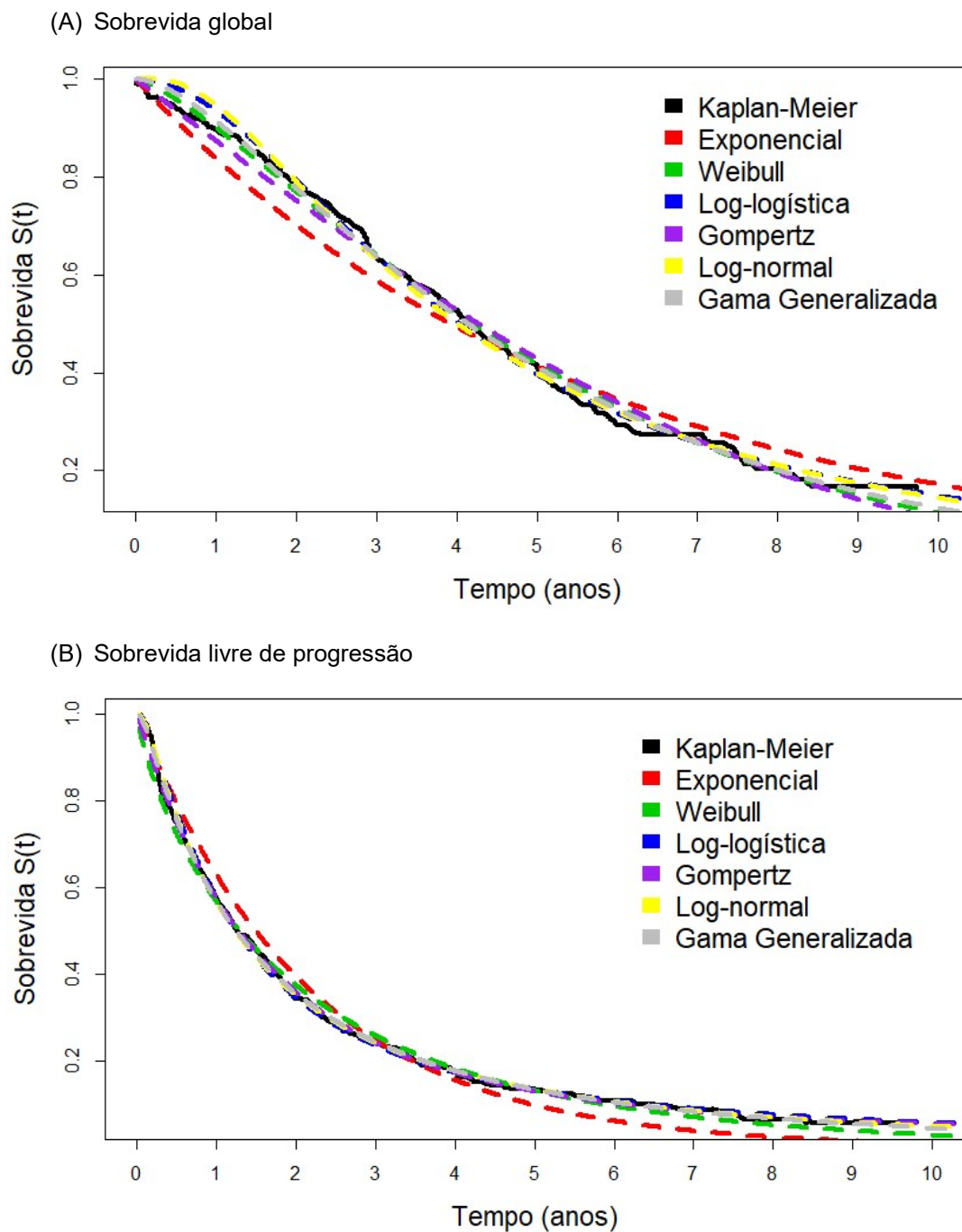
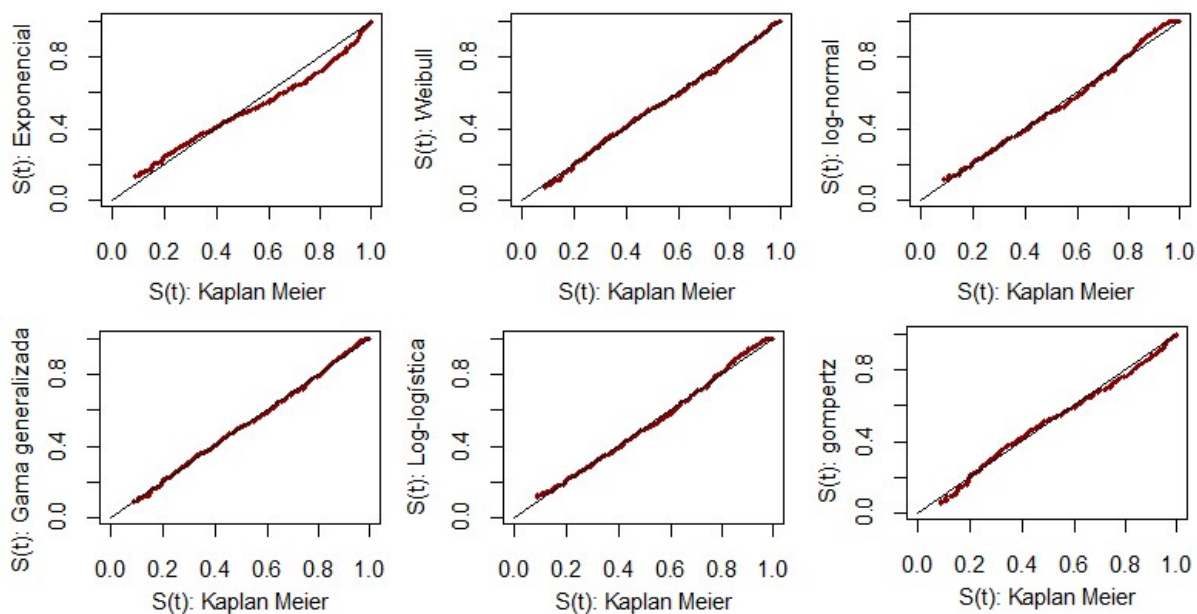


Gráfico 6 – Curvas da Kaplan-Meier obtidas dos dados agregados para o Anastrozol+Fulvestranto vs distribuições estimadas – sobrevida global e livre de progressão

Para os dois braços de tratamentos as curvas estimadas que melhor se ajustaram visualmente aos dados foram as distribuições Weibull e gama generalizada comparado com a curva de Kaplan-Meier para a sobrevida global(A). Para a sobrevida livre de progressão(B) visualmente todas as curvas estão bem próximas sendo assim possível descartar apenas as distribuições exponencial e Weibull para ambos os

braços de tratamento, conforme o Gráfico 5 e Gráfico 6. Os parâmetros calculados das distribuições estão disponibilizados no APÊNDICE F.

(A) Sobrevida global



(B) Sobrevida livre de progressão

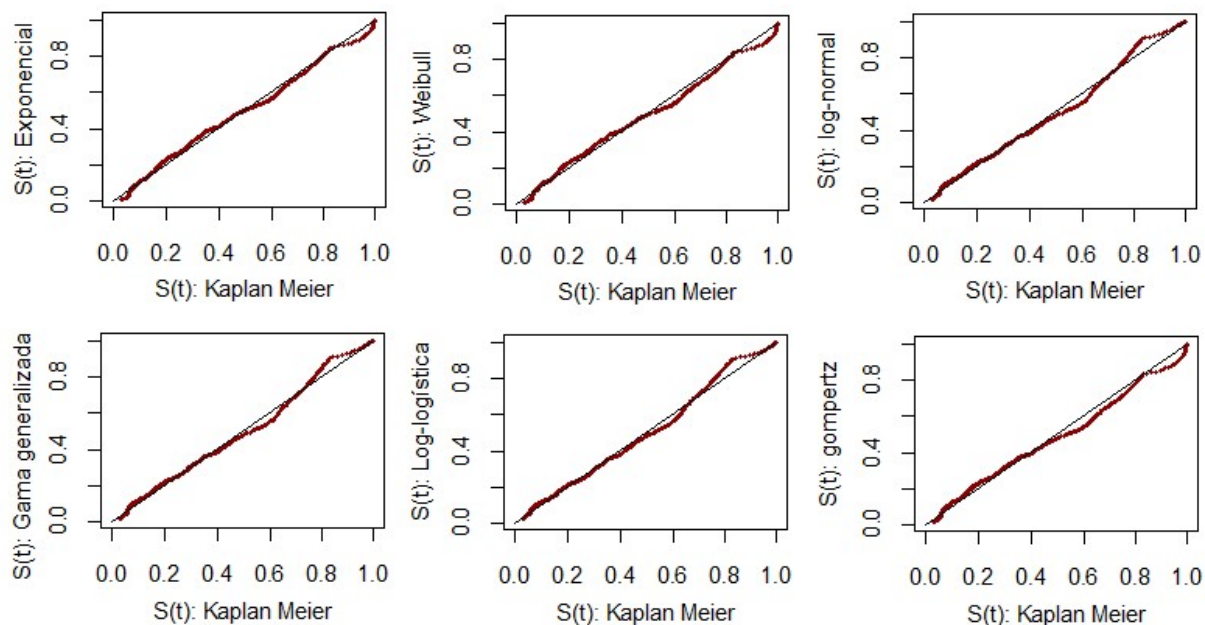


Figura 12 – Curva de sobrevivência de Kaplan-Meier versus curvas de sobrevivência estimadas propostas – tratamento Anastrozol, Sobrevida global e Livre de progressão

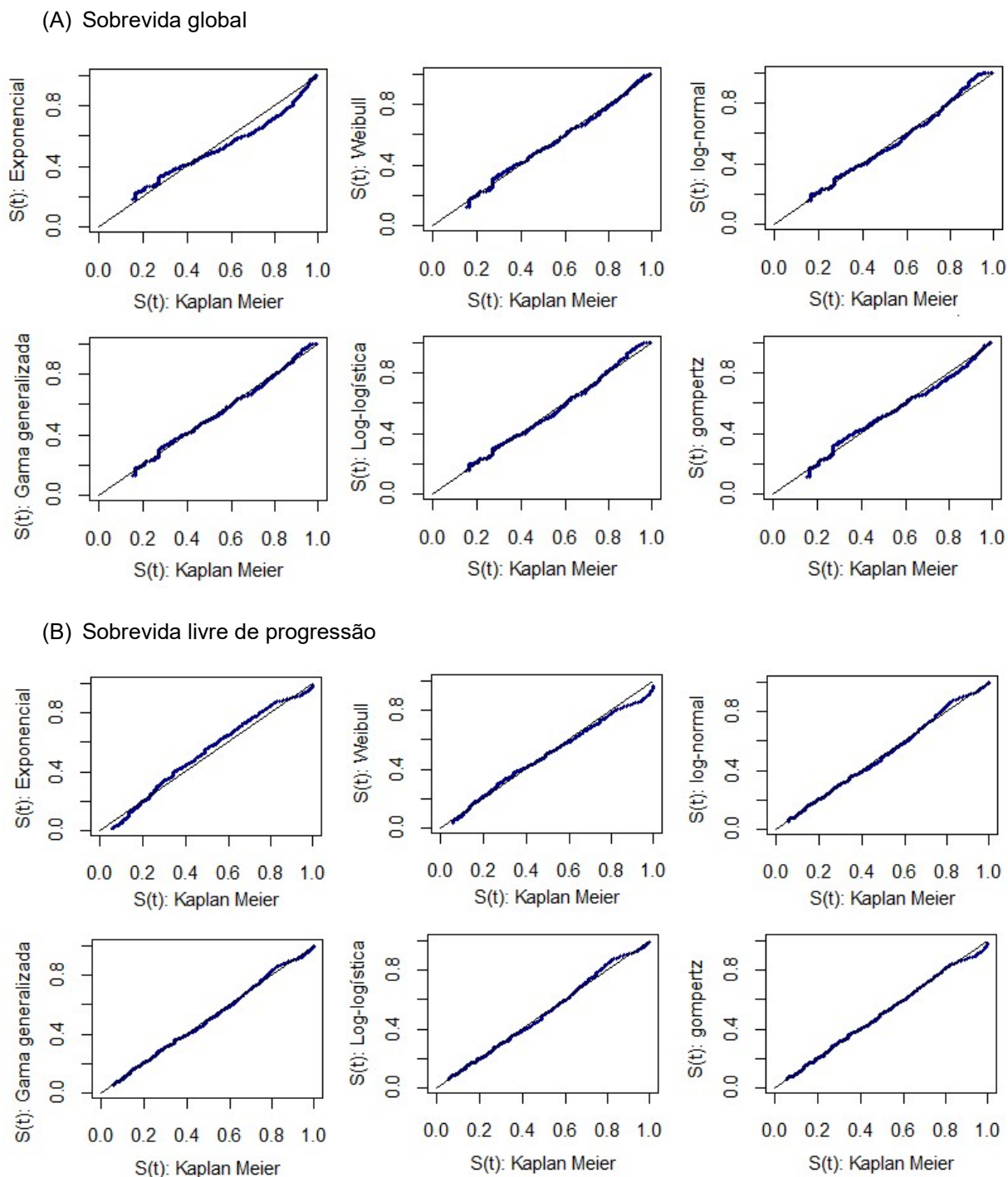


Figura 13 – Curva de sobrevivência de Kaplan-Meier versus curvas de sobrevivência estimadas propostas – tratamento Anastrozol+Fulvestranto, Sobrevida global e livre de progressão

Na inspeção visual baseada na análise dos dados agregados verifica-se que em ambos os braços de tratamento as curvas estimadas Weibull e gama generalizada são as mais adequadas aos dados do estudo para a sobrevida global(A), já no caso da sobrevida livre de progressão(B) para os dois tratamentos visualmente não se

consegue descartar nenhuma das curvas, todas parecem ter um bom ajuste, conforme Figura 12 e Figura 13.

Ajuste dos modelos

Utilizando os critérios de informação AIC e BIC verificou-se que a distribuição gama generalizada possuiu o menor valor para os dois testes e para os dois braços de tratamento, sendo a curva de Weibull a segunda opção no caso da Sobrevida global.

Para a sobrevida livre de progressão a curva log-logística foi a primeira opção para o tratamento Anastrozol e a segunda opção a log-normal. Para Anastrozol+Fulvestranto foram gama generalizada e Gompertz, respectivamente, conforme resultados demonstrados na Tabela 2.

Tabela 2: Estatísticas do critério de informação AIC e BIC, sobrevida global e livre de progressão, para os tratamentos Anastrozol e Anastrozol+Fulvestranto, para os dados agregados

Distribuição	Sobrevida Global				Sobrevida livre de progressão			
	Anastrozol		Anastrozol+Fulvestranto		Anastrozol		Anastrozol+Fulvestranto	
	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
Exponencial	-2903.829	-2894.301	-3156.700	-3146.931	-4431.018	-4421.212	-3514.812	-3505.074
Weibull	-5153.568	-5139.276	-5088.019	-5073.365	-4458.237	-4443.529	-4908.973	-4894.366
Log-normal	-4531.393	-4517.101	-4664.436	-4649.782	-5319.177	-5304.469	-6501.675	-6487.067
Gama Generalizada	-5526.168	-5507.113	-5229.058	-5209.520	-5318.292	-5298.681	-6668.986	-6649.510
Log-logística	-4825.909	-4811.617	-4986.192	-4971.538	-5622.003	-5607.295	-6135.980	-6121.373
Gompertz	-4137.853	-4123.561	-4379.635	-4364.981	-4837.604	-4822.896	-6633.494	-6618.887

Nota: AIC: *Akaike's Information Criterion* e BIC: *Bayesian Information Criterion*

Comparação das curvas escolhidas, dados individualizados e agregados

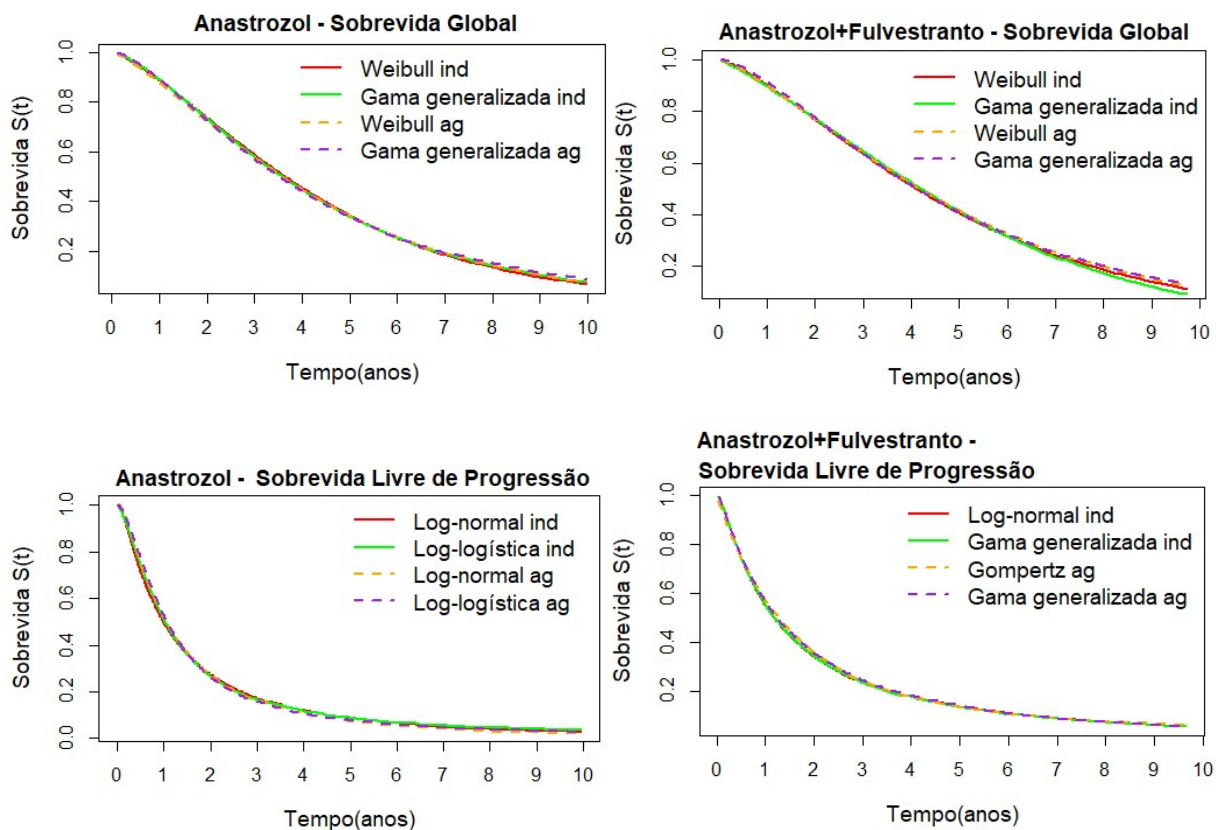


Gráfico 7 – Comparação das curvas com dados individuais vs agregados, sobrevida global e sobrevida livre de progressão

Analisando visualmente os resultados dos dados individualizados versus agregados verifica-se que na prática não existem grandes diferenças entre as curvas de sobrevivência estimadas de dados individuais das curvas dos dados agregados. Tanto para ambos os braços de tratamento, quanto para os desfechos de sobrevida, conforme Gráfico 7 e APÊNDICE C.

5 DISCUSSÃO

5.1 Sobre as diretrizes

Após a recuperação de mais de 300 documentos na base Medline, agências de avaliação de tecnologias em saúde (ATS) e contato com a *International Network of Agencies for Health Technology Assessment* (INAHTA), foi possível identificar os principais documentos disponíveis que continham recomendações explícitas sobre extrapolação de dados. Apesar do grande número de resultados iniciais, ressalta-se que apenas 3 documentos foram incluídos no trabalho por satisfazerem os critérios de incluir de forma explícita alguma recomendação sobre extrapolação

Os documentos encontrados, oriundos de três agências distintas de ATS (PBAC, na Austrália; NICE, na Inglaterra e CONITEC, no Brasil), citaram em consenso as distribuições: exponencial, Weibull, log-normal e Gompertz. Adicionalmente, apesar de não haver consenso, foram também citadas em alguns dos documentos as curvas de distribuições log-logística, gama e gama-generalizada. Dentre as diferenças nos documentos em relação ao uso das distribuições, tem-se que no PBAC foi recomendado as distribuições log-logística e gama, enquanto o NICE usou log-logística e a gama generalizada. Para avaliar o ajuste das curvas extrapoladas, as três diretrizes citaram os critérios de: Inspeção visual, onde deve-se comparar os modelos que mais se ajustam graficamente a curva de Kaplan-Meier e os critérios de informação de *Akaike* (AIC) e Critério de informação *bayesiano* (BIC), onde a escolha do modelo é feita pelo menor valor obtido. Desta forma, seguindo as recomendações mais recentes (propostas pelo documento do NICE), pode-se identificar a necessidade de minimamente utilizar as cinco principais distribuições (exponencial, Weibull, log-normal, Gompertz e gama generalizada), minimizando um possível viés de seleção. Iguamente, para a avaliação do ajuste das curvas extrapoladas, observa-se consenso nos três critérios para a comparação das extrapolações.

A partir da busca estruturada aqui conduzida, é possível observar que a extrapolação de dados é um campo com diretrizes ainda limitadas no campo da ATS, precisando ser mais discutido. Diferentemente de outros critérios, como por exemplo, a análise de sensibilidade, um campo já com um alto grau de consenso sobre as estratégias metodológicas na análise de incertezas paramétricas[26], não é possível afirmar que a extrapolação de dados com curvas de sobrevivência tenha a mesma consistência nas diretrizes consultadas. Contudo, conforme a evolução das técnicas

estatísticas e avanço na medicina oncológica, é muito provável que as extrapolações sejam cada vez mais exploradas nas avaliações econômicas e esse seja um campo a ser aprimorado em breve como ressaltam discussões recentes[27–29].

Apesar de suas fortalezas, o presente levantamento possui algumas limitações. Mesmo sendo a revisão rápida bem aplicada, ela tem suas limitações inerentes a amplitude restrita de fontes de busca, podendo não conseguir identificar alguns documentos. Com o escopo também restrito, destaca-se que a busca foi focada em diretrizes, não havendo espaço para inclusão de artigos que discutissem os métodos para extrapolação e que igualmente apresentam pontos importantes em relação às escolhas metodológicas. Apesar disso, é pouco provável que tais documentos sejam divergentes dos pontos específicos aqui discutidos em relação às principais distribuições de probabilidade e seus critérios de escolha[30,31].

5.2 Sobre as extrapolações

Graças ao sucesso do contato com autores, mais especificamente Dr. William Barlow (Universidade de Washington -EUA), que prontamente disponibilizou os dados individualizados, o presente projeto teve sucesso em sua condução e objetivos sendo possível realizar a comparação entre as curvas estimadas dos dados extraídos por digitalização das imagens do estudo com os dados individualizados cedidos. Esse é apenas mais um reflexo do quanto os novos paradigmas da ciência aberta podem contribuir para a evolução dos métodos em pesquisa.

Primeiramente, pode-se extrair dos resultados que houve semelhança entre as curvas plotadas dos dados individuais versus dados agregados obtidas da digitalização das imagens, mostrando ser plausível utilizar os dados de imagens digitalizadas na ausência de dados individualizados. Depois de estimadas as curvas paramétricas para os dois conjuntos de dados (individual e agregado), verificou-se que para o desfecho sobrevida global e para ambos os tratamentos poderíamos usar as distribuições de Weibull ou gama generalizada, com pouca diferença entre as duas, para fazer a extrapolação com os dados obtidos das imagens. Já na sobrevida livre de progressão as curvas ficaram quase todas bem próximas da curva de Kaplan-Meier para os dois conjuntos de dados e braços de tratamento, ficando difícil a escolha por inspeção visual, restando o critério de avaliação para a escolha das curvas, ficando

entre log-normal e log-logística para o tratamento com anastrozol, Gompertz e gama generalizada para anastrozol+ fulvestranto.

Como uma de suas fortalezas o presente estudo adota um método robusto, criterioso e transparente de execução. Embora já bastante difundida, a extrapolação com curvas de sobrevivência ainda tem suas estratégias metodológicas voltadas ao nível de dados individualizados. Essa é uma importante barreira dado que o acesso a esse tipo de dado ainda é um desafio. Em publicações recentes no tema de acesso a dados individualizados, Scutt (2020) [32] e van Middelkoop (2020) [33] discutem a dificuldade de realizar estudos com dados individualizados de pacientes, já que muitos pesquisadores ainda não compartilham os dados de seus estudos. Ao nosso conhecimento, esse é o primeiro estudo que compara os métodos de extrapolação baseados em ambos os dados individualizados e agregados. Não exatamente neste ponto, outras pesquisas tem buscado analisar a incerteza das extrapolações por meio de sua comparação com dados reais. Como exemplo, Ash Bullement et al. (2019) [34], após revisar 11 estudos de avaliações de tecnologias concluíram que as extrapolações são importantes, pois seus resultados a longo prazo irão convergir com as publicações posteriores. De forma semelhante, Jodi Gray et al 2020 compararam a sobrevivência obtida com as curvas paramétricas padrão e modelos mais flexíveis com os dados de uma coorte de 15 grupos por tipo de câncer e faixa-etária, chegando à conclusão de que até mesmo modelos mais flexíveis como spline tiveram um ajuste melhor aos dados, devendo ser incluídos nos modelos para extrapolar dados de sobrevivência do câncer [35].

Como limitações deste estudo, destaca-se que os resultados aqui obtidos, foram limitados a uma condição e tecnologia na área de oncologia. Todavia, ressalta-se não impedimentos para que os métodos e códigos aqui elaborados e executados possam ser replicados em qualquer contexto de análise de curvas, desde que de posse dos dados agregados ou suas estimativas. Assim, é importante que as condutas aqui apresentadas sejam também testadas em outros contextos e com outras tecnologias. Apesar do consenso nas curvas de sobrevida global, alguma inconsistência foi observada nas curvas de sobrevida livre de progressão ao comparar os métodos pela estimativa agregada versus individual. Contudo, é importante considerar que havia muita incerteza na escolha dessas curvas nos dados individualizados, dada a proximidade dos valores de AIC e BIC, assim como sua

inspeção visual. Por fim, tais pontos apenas ressaltam que a extrapolação com base em dados agregados não deva ser uma escolha de preferência na modelagem em ATS, mas uma opção factível ao contexto de impossibilidade de acesso de dados individualizados.

6 CONCLUSÕES

As principais diretrizes em ATS são ainda escassas e pouco claras em respeito à aplicação e condução da extrapolação de dados com curvas de sobrevivência. Dentro dos pontos de consenso, observa-se a necessidade de minimamente escolher mais de uma curva, sendo as principais: exponencial, Weibull, log-logística, log-normal, gama generalizada e Gompertz. Deve-se verificar as curvas por inspeção visual para descartar as distribuições que não se ajustam aos dados e adotar testes de critérios de informação (AIC/BIC) para confirmar quais dentre as curvas estimadas estão mais aderentes aos dados.

Após a estimativa de curvas de sobrevivência a partir de dados agregados, todos os principais métodos de seleção de curvas adotados nos modelos com dados individualizados puderam ser executados com dados agregados. A extrapolação aqui testada pode ser avaliada por meio da inspeção visual e testes de critério de informação (AIC/BIC) igualmente aos dados individualizados.

A necessidade de dados individualizados é ainda uma barreira para a implementação de modelos de extrapolação de dados com curvas de sobrevivência. Modelos de ajuste baseados em dados agregados são uma alternativa e, a partir de uma comparação baseada em dados reais, tais modelos se mostraram válidos e factíveis a este contexto.

Mesmo sendo uma alternativa factível, as conclusões aqui alcançadas são baseadas em um contexto isolado e específico. Mais estudos devem ser conduzidos com outros contextos e terapias buscando ampliar o escopo de aplicação.

REFERÊNCIAS

1. Piola SF, Vieira FS. Consolidação do gasto com ações e serviços públicos de saúde: trajetória e percalços no período de 2003 a 2017. 2017;
2. Araújo JD de. Polarização epidemiológica no Brasil. *Epidemiol e Serviços Saúde*. 2012;21:533–8.
3. Saúde B-M da. Política Nacional de Gestão de Tecnologias em Saúde. [Série B. Textos Básicos em Saúde]. 2011.
4. Novaes HMD, de Soárez PC. Organizações de avaliação de tecnologias em saúde (ATS): Dimensões do arcabouço institucional e político. *Cad Saude Publica*. 2016;32:1–14.
5. Saúde M da. PORTARIA Nº 3.323 DE 27 DE DEZEMBRO DE 2006 [Internet]. 2006 [cited 2020 May 3]. Available from: https://bvsms.saude.gov.br/bvs/saudelegis/gm/2006/prt3323_27_12_2006.html
6. Brasil - Ministério da Saúde. Portaria nº. 2.915, de 12 de Dezembro de 2011. Ministério da Saúde [Internet]. 2011;7. Available from: http://bvsms.saude.gov.br/bvs/saudelegis/gm/2011/prt2915_12_12_2011.html
7. Brasil. Lei nº 12.401, de 28 de abril de 2011. Altera a Lei no 8.080, de 19 de setembro de 1990, para dispor sobre a assistência terapêutica e a incorporação de tecnologia em saúde no âmbito do Sistema Único de Saúde - SUS. *Diário Of da União* [Internet]. 2011;1–3. Available from: http://www.planalto.gov.br/ccivil_03/_Ato2004-2006/2006/Lei/L11265.htm.
8. Brasil. Entendendo a Incorporação de Tecnologias em Saúde no SUS : como se envolver. [Internet]. 2016. Available from: www.htai.org
9. Brasil. Ministério da Saúde. Secretaria de Ciência T e IED de C e T. Diretrizes metodológicas : Diretriz de Avaliação Econômica [Internet]. 2. ed. Brasília; 2014. Available from: http://conitec.gov.br/images/Artigos_Publicacoes/Diretrizes/AVE.pdf
10. Drummond FM, Mark JS, Karl C, Greg LS, George WT. *Methods for Economic Evaluation of Health Care Programmes*. Oxford University Press; 2015.
11. Briggs A, Claxton K, Mark JS. *Decision Modelling for Health Economic Evaluation*. Published in the United States by Oxford University Press Inc., New York ©; 2006.
12. Colosimo EA, Giolo SR. *Análise De Sobrevida Aplicada*. 1ª edição. São Paulo: Blucher; 2006.
13. Faraway JJ, Tanner M, Zidek J, Abramovich F, Ritov Y, Afifi A, et al. *Modelling Survival Data in Medical Research Modelling Survival Data in Medical Research Third*

Edition Practical Multivariate Analysis, Fifth Edition Interpreting Data: A First Course in Statistics Introduction to Probability with R K. Baclawski Linear Alg. 2015.

14. Carvalho, Marília Sá; Andreozzi, Valeska Lima; Codeço, Claudia Torres; Campos, Dayse Pereira; Barbosa, Maria Tereza Serrano; Shimakura SE. Análise de sobrevivência. 2^a. Fiocruz, editor. Rio de Janeiro; 2011.

15. Caetano R, da Silva RM, Pedro ÉM, de Oliveira IAG, Biz AN, Santana P. Incorporation of new medicines by the national commission for incorporation of technologies, 2012 to June 2016. Cienc e Saude Coletiva. 2017;22:2513–25.

16. National Cancer Institute at the National Institutes of Health [Internet]. [cited 2020 May 4]. Available from: <https://www.cancer.gov/publications/dictionaries/cancer-terms/def/>

17. Latimer N. NICE DSU technical support document 14: survival analysis for economic evaluations alongside clinical trials-extrapolation with patient-level data. Sheff Rep by Decis Support Unit [Internet]. 2011; Available from: http://www.nicedsu.org.uk/NICE_DSU_TSD_Survival_analysis.updated_March_2013.v2.pdf

18. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68:394–424.

19. Instituto Nacional de Câncer José Alencar Gomes da Silva. Estimativa 2020: incidência de câncer no Brasil [Internet]. Rio de Janeiro; 2019. Available from: <http://www.inca.gov.br>

20. Conitec CN de I de T do S. Diretrizes Diagnósticas e Terapêuticas do Carcinoma de Mama. 2018;38.

21. Tricco AC, Langlois EV SS. Rapid reviews to strengthen health policy and systems: a practical guide. Mon Not R Astron Soc [Internet]. 2017;385:453–67. Available from: <https://apps.who.int/iris/bitstream/handle/10665/258698/9789241512763-eng.pdf?sequence=1>

22. Mehta RS, Barlow WE, Albain KS, Vandenberg TA, Dakhil SR, Tirumali NR, et al. Combination anastrozole and fulvestrant in metastatic breast cancer. N Engl J Med. 2012;367:435–44.

23. Mehta RS, Barlow WE, Albain KS, Vandenberg TA, Dakhil SR, Tirumali NR, et al. Overall Survival with Fulvestrant plus Anastrozole in Metastatic Breast Cancer. N Engl J Med. 2019;380:1226–34.

24. Rohatgi A. WebPlotDigitizer [Internet]. San Francisco, California, USA; Available from: <https://automeris.io/WebPlotDigitizer>

25. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ WV, editor. *Cochrane Handbook for Systematic Reviews of Interventions* [Internet]. 2019. Available from: www.training.cochrane.org/handbook
26. Briggs AH, Weinstein MC, Fenwick EAL, Karnon J, Sculpher MJ, Paltiel AD. Model parameter estimation and uncertainty analysis: A report of the ISPOR-SMDM modeling good research practices task force working group-6. *Med Decis Mak*. 2012;32:722–32.
27. Guyot P, Ades AE, Beasley M, Lueza B, Pignon J-P, Welton NJ. Extrapolation of Survival Curves from Cancer Trials Using External Information. *Med Decis Mak* [Internet]. 2017;37:353–66. Available from: <http://journals.sagepub.com/doi/10.1177/0272989X16670604>
28. Bullement A, Willis A, Amin A, Schlichting M, Hatswell AJ, Bharmal M. Evaluation of survival extrapolation in immuno-oncology using multiple pre-planned data cuts: learnings to aid in model selection. *BMC Med Res Methodol* [Internet]. 2020;20:103. Available from: <https://bmcmmedresmethodol.biomedcentral.com/articles/10.1186/s12874-020-00997-x>
29. Vickers A. An Evaluation of Survival Curve Extrapolation Techniques Using Long-Term Observational Cancer Data. *Med Decis Mak* [Internet]. 2019;39:926–38. Available from: <http://journals.sagepub.com/doi/10.1177/0272989X19875950>
30. Bojke L, Manca A, Asaria M, Mahon R, Ren S, Palmer S. How to Appropriately Extrapolate Costs and Utilities in Cost-Effectiveness Analysis. *Pharmacoeconomics* [Internet]. 2017;35:767–76. Available from: <http://link.springer.com/10.1007/s40273-017-0512-6>
31. Hawkins N, Grieve R. Extrapolation of Survival Data in Cost-effectiveness Analyses: The Need for Causal Clarity. *Med Decis Mak* [Internet]. 2017;37:337–9. Available from: <http://journals.sagepub.com/doi/10.1177/0272989X17697019>
32. Scutt P, Woodhouse LJ, Montgomery AA, Bath PM. Data sharing: experience of accessing individual patient data from completed randomised controlled trials in vascular and cognitive medicine. *BMJ Open* [Internet]. 2020;10:e038765. Available from: <https://bmjopen.bmj.com/lookup/doi/10.1136/bmjopen-2020-038765>
33. van Middelkoop M, Lohmander S, Bierma-Zeinstra SMA. Sharing data—taming the beast: barriers to meta-analyses of individual patient data (IPD) and solutions. *Br J Sports Med* [Internet]. 2020;54:822–4. Available from: <https://bjsm.bmj.com/lookup/doi/10.1136/bjsports-2019-101892>
34. Bullement A, Meng Y, Cooper M, Lee D, Harding TL, O'Regan C, et al. A review and validation of overall survival extrapolation in health technology assessments of cancer immunotherapy by the National Institute for Health and Care Excellence: how did the initial best estimate compare to trial data subsequently made availabl. *J Med*

Econ [Internet]. Taylor & Francis; 2019;22:205–14. Available from: <https://doi.org/10.1080/13696998.2018.1547303>

35. This R, Attribution-noncommercial-noderivs CC, By-nc-nd CC, If T, Rose W. Extrapolation of survival curves using standard parametric models and flexible parametric spline models: comparisons in large registry cohorts with advanced cancer. 2020; Available from: https://www.researchgate.net/publication/345319991_Extrapolation_of_Survival_Curves_Using_Standard_Parametric_Models_and_Flexible_Parametric_Spline_Models_Comparisons_in_Large_Registry_Cohorts_with_Advanced_Cancer

APÊNDICES

APÊNDICE A – Mensagem enviada a todas as agências da INAHTA, solicitando informações sobre diretrizes de extrapolação de dados usando análise de sobrevivência.

Dear Mr./Mrs

I hope this message find you very well.

My name is Solange Borges and I am a HTA Master's Degree student in the National Cardiology Institute of Ministry of Health in Brazil.

I am conducting a research to find specific Guidelines (or specific topics in general Guidelines) about Extrapolating Clinical Research Data Beyond Follow-up Using Survival Analysis among publications of INAHTA Members and I would like to count on your help to know if your institution has any kind of publication like that.

I really appreciate to count on your kind help and collaboration to share this information to insert in my research.

Thank you very much in advance and best regards,

Solange Borges

HTA Specialist – Brazil

APÊNDICE B – Formulário de extração

Detalhes do documento	
Título:	
Autor:	
Ano de publicação:	
Objetivo (s):	
Tipo de Estudo:	
País:	
Contexto:	
Resultados extraídos (recomendações)	

APÊNDICE C – Comparação das curvas com dados individuais vs agregados, sobrevida global e sobrevida livre de progressão

Anastrozol - Sobrevida Global					Anastrozol+ Fulvestranto - Sobrevida Global				
tempo	Weibull_ind	Gamgen_ind	Weibull_ag	Gamgen_ag	tempo	Weibull_ind	Gamgen_ind	Weibull_agr	Gamgen_agr
0,12	0,993	0,995	0,992	0,996	0,04	0,999	0,998	0,999	1,000
0,15	0,991	0,993	0,989	0,995	0,07	0,997	0,995	0,997	0,999
0,18	0,989	0,991	0,986	0,993	0,11	0,995	0,992	0,994	0,998
0,19	0,988	0,990	0,985	0,992	0,13	0,993	0,990	0,993	0,997
0,21	0,986	0,989	0,983	0,991	0,15	0,992	0,989	0,991	0,996
0,23	0,984	0,987	0,981	0,989	0,16	0,991	0,988	0,991	0,995
0,25	0,982	0,985	0,979	0,988	0,17	0,990	0,987	0,990	0,995
0,29	0,978	0,982	0,974	0,984	0,21	0,987	0,983	0,987	0,993
0,31	0,976	0,980	0,972	0,982	0,27	0,982	0,977	0,982	0,989
0,32	0,975	0,979	0,971	0,981	0,3	0,980	0,974	0,979	0,987
0,43	0,963	0,967	0,957	0,970	0,38	0,972	0,966	0,971	0,981
0,50	0,955	0,959	0,948	0,962	0,42	0,968	0,962	0,967	0,978
0,51	0,954	0,958	0,947	0,961	0,5	0,960	0,954	0,959	0,970
0,56	0,948	0,952	0,940	0,954	0,52	0,958	0,951	0,957	0,969
0,61	0,941	0,945	0,933	0,948	0,53	0,957	0,950	0,956	0,968
0,63	0,939	0,943	0,930	0,945	0,6	0,949	0,943	0,948	0,961
0,64	0,937	0,941	0,929	0,944	0,63	0,946	0,939	0,945	0,957
0,68	0,932	0,936	0,923	0,938	0,66	0,942	0,936	0,941	0,954
0,71	0,928	0,932	0,919	0,934	0,72	0,936	0,929	0,935	0,948
0,72	0,927	0,931	0,918	0,933	0,73	0,934	0,928	0,933	0,947
0,73	0,926	0,929	0,916	0,931	0,84	0,921	0,915	0,920	0,934
0,77	0,920	0,924	0,910	0,925	0,85	0,920	0,914	0,919	0,933
0,78	0,919	0,922	0,909	0,924	0,91	0,913	0,907	0,912	0,925
0,79	0,917	0,921	0,908	0,923	0,93	0,910	0,905	0,909	0,923
0,82	0,913	0,917	0,903	0,918	0,97	0,905	0,900	0,905	0,918
0,83	0,912	0,915	0,902	0,917	1,02	0,899	0,894	0,898	0,912
0,85	0,909	0,913	0,899	0,914	1,04	0,897	0,892	0,896	0,909
0,91	0,901	0,904	0,890	0,905	1,05	0,895	0,891	0,895	0,908
0,95	0,895	0,898	0,884	0,898	1,07	0,893	0,888	0,892	0,905
0,96	0,894	0,896	0,882	0,897	1,14	0,884	0,880	0,883	0,896
0,99	0,889	0,892	0,878	0,892	1,29	0,864	0,862	0,864	0,876
1,01	0,886	0,889	0,875	0,889	1,3	0,863	0,860	0,863	0,875
1,02	0,885	0,888	0,873	0,887	1,32	0,861	0,858	0,860	0,872
1,03	0,883	0,886	0,872	0,886	1,33	0,859	0,857	0,859	0,871
1,04	0,882	0,885	0,870	0,884	1,35	0,857	0,854	0,856	0,868
1,07	0,878	0,880	0,866	0,879	1,37	0,854	0,852	0,854	0,865
1,12	0,870	0,872	0,858	0,871	1,38	0,853	0,850	0,852	0,864
1,13	0,869	0,871	0,857	0,870	1,39	0,851	0,849	0,851	0,862
1,15	0,866	0,868	0,854	0,867	1,43	0,846	0,844	0,846	0,857
1,16	0,864	0,866	0,852	0,865	1,44	0,845	0,843	0,844	0,856
1,17	0,863	0,865	0,851	0,863	1,48	0,839	0,838	0,839	0,850
1,18	0,861	0,863	0,849	0,862	1,5	0,837	0,836	0,837	0,847
1,19	0,860	0,862	0,847	0,860	1,53	0,833	0,832	0,833	0,843
1,30	0,843	0,844	0,830	0,842	1,58	0,826	0,826	0,826	0,836
1,32	0,840	0,841	0,827	0,839	1,61	0,822	0,822	0,822	0,832
...
9,98	0,063	0,072	0,071	0,083	9,74	0,112	0,091	0,122	0,131

Anastrozol - Sobrevida livre de progressão					Anastrozol+Fulvestranto - Sobrevida livre de progressão				
tempo	Lnormal_ind	Llogist_ind	Lnormal_agr	Llogist_agr	tempo	Lnormal_ind	Gamgen_ind	Gompertz_agr	Gamgen_agr
0,01	1,000	0,999	1,000	1,000	0,04	0,996	0,995	0,976	0,994
0,03	0,999	0,995	1,000	0,997	0,06	0,990	0,989	0,965	0,986
0,05	0,995	0,990	0,998	0,994	0,07	0,986	0,985	0,959	0,982
0,06	0,992	0,987	0,997	0,992	0,08	0,982	0,981	0,953	0,978
0,07	0,989	0,983	0,995	0,989	0,1	0,973	0,972	0,942	0,968
0,09	0,980	0,976	0,991	0,984	0,11	0,968	0,967	0,936	0,963
0,10	0,975	0,971	0,988	0,981	0,13	0,957	0,956	0,925	0,953
0,12	0,965	0,963	0,981	0,974	0,14	0,952	0,951	0,920	0,947
0,13	0,959	0,958	0,978	0,971	0,16	0,940	0,939	0,909	0,936
0,14	0,954	0,953	0,974	0,967	0,17	0,934	0,934	0,904	0,931
0,15	0,947	0,949	0,969	0,963	0,18	0,929	0,928	0,899	0,925
0,16	0,941	0,944	0,965	0,959	0,19	0,923	0,922	0,893	0,920
0,17	0,935	0,939	0,960	0,955	0,2	0,917	0,916	0,888	0,914
0,18	0,929	0,933	0,955	0,951	0,21	0,911	0,910	0,883	0,908
0,19	0,922	0,928	0,950	0,946	0,22	0,905	0,904	0,878	0,903
0,20	0,915	0,923	0,945	0,942	0,23	0,899	0,898	0,873	0,897
0,21	0,909	0,917	0,940	0,937	0,24	0,893	0,892	0,868	0,891
0,22	0,902	0,912	0,934	0,932	0,25	0,887	0,886	0,863	0,886
0,23	0,895	0,906	0,928	0,928	0,26	0,881	0,881	0,858	0,880
0,24	0,888	0,901	0,923	0,923	0,27	0,875	0,875	0,853	0,875
0,25	0,882	0,895	0,917	0,918	0,28	0,869	0,869	0,848	0,869
0,26	0,875	0,889	0,911	0,913	0,29	0,863	0,863	0,843	0,863
0,27	0,868	0,884	0,905	0,908	0,3	0,857	0,857	0,838	0,858
0,28	0,861	0,878	0,899	0,903	0,33	0,840	0,840	0,824	0,842
0,29	0,855	0,872	0,893	0,897	0,34	0,834	0,834	0,819	0,836
0,31	0,841	0,860	0,881	0,887	0,36	0,823	0,823	0,810	0,826
0,32	0,834	0,855	0,875	0,881	0,38	0,811	0,812	0,801	0,815
0,33	0,828	0,849	0,868	0,876	0,39	0,806	0,807	0,797	0,810
0,36	0,808	0,831	0,850	0,859	0,4	0,801	0,801	0,792	0,805
0,37	0,802	0,825	0,844	0,854	0,41	0,795	0,796	0,788	0,800
0,38	0,795	0,819	0,837	0,848	0,44	0,779	0,780	0,774	0,785
0,40	0,783	0,808	0,825	0,837	0,45	0,774	0,775	0,770	0,780
0,41	0,776	0,802	0,819	0,831	0,46	0,769	0,770	0,766	0,775
0,42	0,770	0,796	0,813	0,826	0,47	0,764	0,765	0,762	0,770
0,43	0,764	0,790	0,807	0,820	0,5	0,748	0,750	0,749	0,756
0,44	0,758	0,784	0,801	0,814	0,52	0,739	0,740	0,741	0,747
0,46	0,746	0,773	0,788	0,803	0,53	0,734	0,735	0,737	0,742
0,48	0,734	0,761	0,776	0,792	0,54	0,729	0,731	0,733	0,738
0,49	0,728	0,755	0,771	0,786	0,55	0,724	0,726	0,729	0,733
0,50	0,722	0,750	0,765	0,780	0,56	0,720	0,721	0,725	0,729
0,51	0,717	0,744	0,759	0,775	0,59	0,706	0,708	0,713	0,716
0,52	0,711	0,738	0,753	0,769	0,61	0,697	0,699	0,705	0,707
0,53	0,705	0,733	0,747	0,763	0,63	0,688	0,690	0,697	0,699
0,55	0,694	0,722	0,736	0,752	0,64	0,684	0,686	0,694	0,695
0,56	0,689	0,716	0,730	0,746	0,68	0,667	0,669	0,679	0,678
0,59	0,673	0,700	0,713	0,730	0,7	0,658	0,661	0,672	0,671
0,61	0,662	0,689	0,702	0,719	0,71	0,654	0,656	0,668	0,667
0,62	0,657	0,684	0,697	0,713	0,72	0,650	0,652	0,664	0,663
0,63	0,652	0,679	0,692	0,708	0,73	0,646	0,649	0,661	0,659
0,64	0,647	0,674	0,686	0,702	0,75	0,638	0,641	0,654	0,651
0,65	0,642	0,668	0,681	0,697	0,77	0,631	0,633	0,647	0,644
0,67	0,633	0,658	0,671	0,686	0,78	0,627	0,629	0,644	0,640
0,68	0,628	0,653	0,665	0,681	0,79	0,623	0,626	0,640	0,637
0,69	0,623	0,648	0,660	0,676	0,8	0,619	0,622	0,637	0,633
0,7	0,618	0,643	0,655	0,670	0,81	0,616	0,618	0,634	0,630
0,71	0,614	0,638	0,650	0,665	0,84	0,605	0,607	0,624	0,619
0,72	0,609	0,633	0,645	0,660	0,85	0,601	0,604	0,620	0,616
0,73	0,605	0,628	0,640	0,655	0,92	0,577	0,580	0,598	0,592
0,74	0,600	0,624	0,636	0,650	0,93	0,574	0,577	0,595	0,589
0,77	0,587	0,609	0,621	0,635	0,94	0,571	0,573	0,592	0,586
0,79	0,578	0,600	0,612	0,625	0,96	0,564	0,567	0,586	0,580
...
9,95	0,024	0,032	0,016	0,025	9,95	0,051	0,049	0,056	0,049

APÊNDICE D – Implementação dos códigos na linguagem R

```
#####
#Fitting Survival Models to Aggregate Data
#Authors: Solange Borges (solangeborges@gmail.com), Ivan Zimmermann (ivanzricardo@gmail.com)
#Date: 04/13/2021
#####

#Install and load packages
install.packages("readr")
install.packages("flexsurv")
library(readr)
library(flexsurv)

#Load life table data with time and aggregated survival probabilities columns.
data <- read_delim('INSERT CSV FILE PATH',
  ":", escape_double = FALSE, col_types = cols(survival = col_number()),
  locale = locale(decimal_mark = ","),
  trim_ws = TRUE)

  #This .csv can be generated with digitization methods
  #It is importante to the first row as time = 0 and survival = 1

time <- data$time
survival <- data$survival
horizon <- 30 #Extrapolating horizon (time beyond available data)

#Fit the desired functions with non-linear least square method (NLS)
reg_exp <- nls(survival ~ exp(-g_e*time),data, start=list(g_e=0.1)) #Exponential
reg_wei <- nls(survival ~ exp(-l_w*(time^g_w)),data, start=list(l_w=0.1,g_w=0.01)) #Weibull
reg_logl <- nls(survival ~ 1/(1 + l_l*time^g_l),data, start=list(l_l=0.1,g_l=0.01)) #Log-Logistic
reg_gomp <- nls(survival ~ exp(-(g_g)/l_g)*(exp(l_g*time) - 1),data, start=list(l_g=0.1,g_g=0.01),
  algorithm="port", lower=c(-Inf,-Inf)) #Gompertz
reg_lnorm <- nls(survival ~ 1- pnorm(((log(time)-mu_ln)/sd_ln),mu_ln,sd_ln),data, start=list(mu_ln=1,sd_ln=1))
#Lognormal
reg_gam_gen <- nls(survival ~ 1 - pgamma(time, mu_g, sigma, Q),data, start=list(sigma=1,mu_g=1, Q=1),
  algorithm="port", lower=c(-Inf,-Inf,-Inf)) #Generalized gamma

#Store the parameters of the fitted functions
Model <- c("Exponential")
p_reg_exp <- data.frame(summary(reg_exp)[["parameters"]])
Parameter <- row.names(p_reg_exp)
p_reg_exp <- data.frame(Model, Parameter, p_reg_exp)

Model <- c("Weibull")
p_reg_wei <- data.frame(summary(reg_wei)[["parameters"]])
Parameter <- row.names(p_reg_wei)
p_reg_wei <- data.frame(Model, Parameter, p_reg_wei)

Model <- c("Log-logistic")
p_reg_logl <- data.frame(summary(reg_logl)[["parameters"]])
Parameter <- row.names(p_reg_logl)
p_reg_logl <- data.frame(Model, Parameter, p_reg_logl)

Model <- c("Gompertz")
p_reg_gomp <- data.frame(summary(reg_gomp)[["parameters"]])
Parameter <- row.names(p_reg_gomp)
p_reg_gomp <- data.frame(Model, Parameter, p_reg_gomp)
```

```

Model <- c("Lognormal")
p_reg_Inorm <- data.frame(summary(reg_Inorm)[["parameters"]])
Parameter <- row.names(p_reg_Inorm)
p_reg_Inorm <- data.frame(Model, Parameter, p_reg_Inorm)

Model <- c("Generalized Gamma")
p_reg_gam_gen <- data.frame(summary(reg_gam_gen)[["parameters"]])
Parameter <- row.names(p_reg_gam_gen)
p_reg_gam_gen <- data.frame(Model, Parameter, p_reg_gam_gen)

Parameters <- rbind(p_reg_exp,p_reg_wei,p_reg_logl,p_reg_gomp,p_reg_Inorm, p_reg_gam_gen)

#Evaluate visual fitting
##Generate predicted data for fitted curves
data$time_pred <- data$time
data[(nrow(data)+1):(nrow(data)+horizon),"time_pred"] <- seq(from = round(as.numeric(data[nrow(data),1])+1),
digits = 0), to = round(as.numeric(data[nrow(data),1]+ horizon), digits = 0), by = 1)
time_pred <-data$time_pred
data$fit_exp <- exp(-p_reg_exp[1,3]*time_pred)
data$fit_wei <- exp(-p_reg_wei[1,3]*(time_pred^p_reg_wei[2,3]))
data$fit_logl <- 1/(1 + p_reg_logl[1,3]*time_pred^p_reg_logl[2,3])
data$fit_gomp <- exp(-((p_reg_gomp[2,3])/p_reg_gomp[1,3])*(exp(p_reg_gomp[1,3]*time_pred) - 1))
data$fit_Inorm <- 1- pnorm(((log(time_pred)-
p_reg_Inorm[1,3])/p_reg_Inorm[2,3]),p_reg_Inorm[1,3],p_reg_Inorm[2,3])
data$fit_gam_gen <- 1 - pgengamma(time_pred, p_reg_gam_gen[2,3], p_reg_gam_gen[1,3],
p_reg_gam_gen[3,3])

##Plot survival observed data and curves
png("survival_plot.png", width=750, height=450)
plot(data$time_pred,data$survival, main="Survival Curves", xlab = "Time", ylab = "Survival", type = "l",
col="black",lty=1,lwd=3)
line_exp <- lines(data$time_pred,data$fit_exp,col="red",lty=2,lwd=2)
line_wei <- lines(data$time_pred,data$fit_wei,col="green3",lty=2,lwd=2)
line_logl <- lines(data$time_pred,data$fit_logl,col="blue",lty=2,lwd=2)
line_gomp <- lines(data$time_pred,data$fit_gomp,col="purple",lty=2,lwd=2)
line_Inorm <- lines(data$time_pred,data$fit_Inorm,col="yellow",lty=2,lwd=2)
line_gam_gen <- lines(data$time_pred,data$fit_gam_gen,col="gray",lty=2,lwd=2)
legend("topright",
      c("Kaplan-Meier","Exponential", "Weibull", "Log-logistic", "Gompertz", "Lognormal", "Generalized Gamma"),
      fill=c("black","red", "green3", "blue", "purple","yellow","gray"),
      border = F,
      cex=0.9, pt.cex = 1,
      bty = "n"
)
dev.off()

#Evaluate fitting with AIC/BIC statistics (the lower, the better)
AIC(reg_exp, reg_wei, reg_logl, reg_gomp, reg_Inorm, reg_gam_gen)
BIC(reg_exp, reg_wei, reg_logl, reg_gomp, reg_Inorm, reg_gam_gen)

#Export results
write.table(data,file = "Survival_Models_data.csv", sep = ",", col.names = T, row.names = F)
write.table(AUC_models,file = "Survival_Models_AUC.csv", sep = ",", col.names = T, row.names = F)
write.table(Parameters,file = "Survival_Models_parameters.csv", sep = ",", col.names = T, row.names = F)

#visual adjustment of the x = y model comparing kaplan-Meier survival versus estimated curves

data<- read.table("Survival_Models_data.csv ",header = T,sep=",")

```

```

time <- data$time
KM <- data$sobrevida
exponential <- data$fit_exp
weibull <- data$fit_wei
loglogistica <- data$fit_logl
gompertz <- data$fit_gomp
lognormal <- data$fit_lnorm
gamageneralizada <- data$fir_gam_gen

plot(KM,exponential,pch=20,xlim = range(c(0,1)), ylim = range(c(0,1)),col="dark blue",xlab = "S(t): Kaplan
Meier",ylab = "S(t): Exponencial",cex=0.7,cex.lab=0.9, cex.axis=0.9)
lines(c(0,1),c(0,1),type = "l",lty=1)

plot(KM,weibull,pch=20,xlim = range(c(0,1)),ylim = range(c(0,1)),col="dark blue",xlab = "S(t): Kaplan Meier",ylab =
"S(t): Weibull",cex=0.7,cex.lab=0.9, cex.axis=0.9)
lines(c(0,1),c(0,1),type = "l",lty=1)

plot(KM,loglogistica,pch=20,xlim = range(c(0,1)),ylim = range(c(0,1)),col="dark blue",xlab = "S(t): Kaplan
Meier",ylab = "S(t): Log-logística",cex=0.7,cex.lab=0.9, cex.axis=0.9)
lines(c(0,1),c(0,1),type = "l",lty=1)

plot(KM,gompertz,pch=20,xlim = range(c(0,1)),ylim = range(c(0,1)),col="dark blue",xlab = "S(t): Kaplan
Meier",ylab = "S(t): gompertz",cex=0.7,cex.lab=0.9, cex.axis=0.9)
lines(c(0,1),c(0,1),type = "l",lty=1)

plot(KM,lognormal,pch=20,xlim = range(c(0,1)),ylim = range(c(0,1)),col="dark blue",xlab = "S(t): Kaplan
Meier",ylab = "S(t): log-normal",cex=0.7,cex.lab=0.9, cex.axis=0.9)
lines(c(0,1),c(0,1),type = "l",lty=1)

plot(KM,gamageneralizada,pch=20,xlim = range(c(0,1)),col="dark blue",ylim = range(c(0,1)),xlab = "S(t): Kaplan
Meier",ylab = "S(t): Gama generalizada",cex=0.7,cex.lab=0.9, cex.axis=0.9)
lines(c(0,1),c(0,1),type = "l",lty=1)

##### THIS THE END #####

#####
#Fitting Survival Models to individualized Data #
#Authors: Solange Borges (solangeborges@gmail.com), Ivan Zimmermann (ivanzricardo@gmail.com) #
#Date: 04/13/2021 #
#####

#Install and load packages
install.packages("stringr")
require(survival)
require(flexsurv)
require(stringr)

#input data with treatment
data<- read.table("INSERT CSV FILE PATH",header = T,sep=";") #'INSERT CSV FILE PATH'

#NON PARAMETRIC ESTIMATE - KAPLAN-MEIER
ekm<-survfit(Surv(time, status)~TRNO,data=data)
time<- ekm$time
skm<-ekm$surv

#plotting the kaplan meier graph
plot(ekm,col=c("dark red","dark blue"),lty=c(1,1),lwd=3, xlab = "Time (year)", ylab = "Survival S(t)",mark.time =
F,axes=F,cex.lab=1.1, cex.axis=1.1)

```

```
legend(8,1,col=c('dark red','dark blue'), legend=c('treatment'),bty="n",lwd=3,
      axis(1,at=c(0,1,2,3,4,5,6,7,8,9,10,11,12)),axis(2,at=c(0.0,0.2,0.4,0.6,0.8,1.0)),box(),cex=1.1)
```

#PARAMETRIC MODELS

```
Exponential <- flexsurvreg(Surv(time, status)~1,data=data, dist="exp")
Weibull <- flexsurvreg(Surv(time, status)~1,data=data, dist="weibull")
Lognormal <- flexsurvreg(Surv(time, status)~1,data=data, dist="lnorm")
Gengamma <- flexsurvreg(Surv(time, status)~1,data=data, dist="gengamma")
Loglogistic <- flexsurvreg(Surv(time, status)~1,data=data, dist="llogis")
Gompertz <- flexsurvreg(Surv(time, status)~1,data=data, dist="gompertz",inits = c(-
1,1/mean(trat_cancer_anas$time)))
```

#GRAPHIC TO COMPARE PARAMETRIC CURVES WITH THE KAPLA-MEIER CURVE

```
plot(ekm,conf.int = F,lty=1,lwd=2,xlab = "Time (year)",ylab = "Survival S(t)", axes=F)
lines(Exponential, lwd=2,lty = 2, col="purple",conf.int = F,col.ci = F)
lines(Weibull, lwd=2,lty = 2, col="dark blue",conf.int = F,col.ci = F)
lines(Lognormal, lwd=2,lty = 2, col="dark green",conf.int = F,col.ci = F)
lines(Gengamma,lwd=2,lty = 2, col="dark red",conf.int = F,col.ci = F)
lines(Loglogistic,lwd=2,lty = 2, col="orange",conf.int = F,col.ci = F,)
lines(Gompertz,lwd=2,lty = 2, col="grey",conf.int = F,col.ci = F,)
legend(9.5,1,lty = c(1,2,2,2,2,2),lwd=2,col=c("black","purple","dark blue","darkgreen","dark
red","orange","grey"),c("Kaplan-Meier","Exponential","Weibull","Log-normal","Generalized Gamma","Log-
logistic","Gompertz"),bty="n",
      axis(1,at=c(0,1,2,3,4,5,6,7,8,9,10)),axis(2,at=c(0.0,0.2,0.4,0.6,0.8,1.0)),box(),cex=1)
```

#ADJUSTED FUNCTIONS

```
s_Exponential<- summary(Exponential)
s_Weibull<-summary(Weibull)
s_Lognormal<-summary(Lognormal)
s_Gengamma<-summary(Gengamma)
s_Loglogistic<-summary(Loglogistic)
s_Gompertz<-summary(Gompertz)
```

```
Model <- as.data.frame(s_Exponential)
colnames(Model)<- c("time","Exponential","ici","uci")
surv_Exponencial<-Model$Exponential
```

```
Model <- as.data.frame(s_Weibull)
colnames(Model)<- c("time","Weibull","ici","uci")
surv_Weibull<- Model$Weibull
```

```
Model <- as.data.frame(s_Lognormal)
colnames(Model)<- c("time","Lognormal","ici","uci")
surv_Lnormal<- Model$Lognormal
```

```
Model <- as.data.frame(s_Gengamma)
colnames(Model)<- c("time","Gengamma","ici","uci")
surv_Gengamma<- Model$Gengamma
```

```
Model <- as.data.frame(s_Loglogistic)
colnames(Model)<- c("time","Llog","ici","uci")
surv_Llogistic<- Model$Llog
```

```
Model <- as.data.frame(s_Gompertz)
colnames(Model)<- c("time","Gom","ici","uci")
surv_Gompertz<- Model$Gom
```

```
#---CREATING A UNIQUE TABLE AND SAVING THE SURVIVAL CURVES ARCHIVE-----
```

```

dados<-
cbind(time,skm,surv_Exponencial,surv_Weibull,surv_Lnormal,surv_Gengamma,surv_Llogistic,surv_Gompertz)
write.table(dados,file = "Survival_Models_data.csv", sep = ",", col.names = T, row.names = F)

#Graphics to compare the estimated curves with the Kaplan-Meier curve straight x=y

plot(skm,surv_Exponencial,pch=20,xlim = range(c(0,1)), ylim = range(c(0,1)),col="dark red",xlab = "S(t): Kaplan
Meier",ylab = "S(t): Exponential",cex=0.7,cex.lab=0.9, cex.axis=0.9)
lines(c(0,1),c(0,1),type = "l",lty=1)

plot(skm,surv_Weibull,pch=20,xlim = range(c(0,1)), ylim = range(c(0,1)),col="dark red",xlab = "S(t): Kaplan
Meier",ylab = "S(t): Weibull",cex=0.7,cex.lab=0.9, cex.axis=0.9)
lines(c(0,1),c(0,1),type = "l",lty=1)

plot(skm,surv_Lnormal,pch=20,xlim = range(c(0,1)), ylim = range(c(0,1)),col="dark red",xlab = "S(t): Kaplan
Meier",ylab = "S(t): Log-normal",cex=0.7,cex.lab=0.9, cex.axis=0.9)
lines(c(0,1),c(0,1),type = "l",lty=1)

plot(skm,surv_Gengamma,pch=20,xlim = range(c(0,1)), ylim = range(c(0,1)),col="dark red",xlab = "S(t): Kaplan
Meier",ylab = "S(t): Generalized Gamma",cex=0.7,cex.lab=0.9, cex.axis=0.9)
lines(c(0,1),c(0,1),type = "l",lty=1)

plot(skm,surv_Llogistic,pch=20,xlim = range(c(0,1)), ylim = range(c(0,1)),col="dark red",xlab = "S(t): Kaplan
Meier",ylab = "S(t): Log-logistic",cex=0.7,cex.lab=0.9, cex.axis=0.9)
lines(c(0,1),c(0,1),type = "l",lty=1)

plot(skm,surv_Gompertz,pch=20,xlim = range(c(0,1)), ylim = range(c(0,1)),col="dark red",xlab = "S(t): Kaplan
Meier",ylab = "S(t): Gompertz",cex=0.7,cex.lab=0.9, cex.axis=0.9)
lines(c(0,1),c(0,1),type = "l",lty=1)

# TESTING INFORMATION CRITERIA TO DISCARD MODELS THAT DO NOT FIT DATA
aic<- AIC(Exponential,Weibull,Lognormal,Gengamma,Loglogistic,Gompertz)
bic<- BIC(Exponential,Weibull,Lognormal,Gengamma,Loglogistic,Gompertz)
aic_bic <- cbind(aic, bic)
##### THIS THE END #####

```

APÊNDICE E – Manual para extração de dados de imagens

Extraindo informações de gráficos

Como usar o WebPlotDigitizer

Onde obter: <https://automeris.io/WebPlotDigitizer/download.html>

Como citar: <https://automeris.io/WebPlotDigitizer/citation.html>

Baixar versões de desktop

WebPlotDigitizer é inteiramente financiado por doações de usuários

[Donate](#)

Pacote	Soma MD5
Windows 64 bits	7cf0e99cd5bd9d440cf5576d1da96aa6c
Windows 32 bits	f9f4df1f8ece4aafca159d2be518291
MacOS 64 bits	b5ef8304cbd0e50aa77d126f7d15607c
Linux 64 bits	ed8fca1fbb24e24658bbd4966ee985
Módulo NodeJS	4f6fc89f93800e4f09dd31650b8a2dea

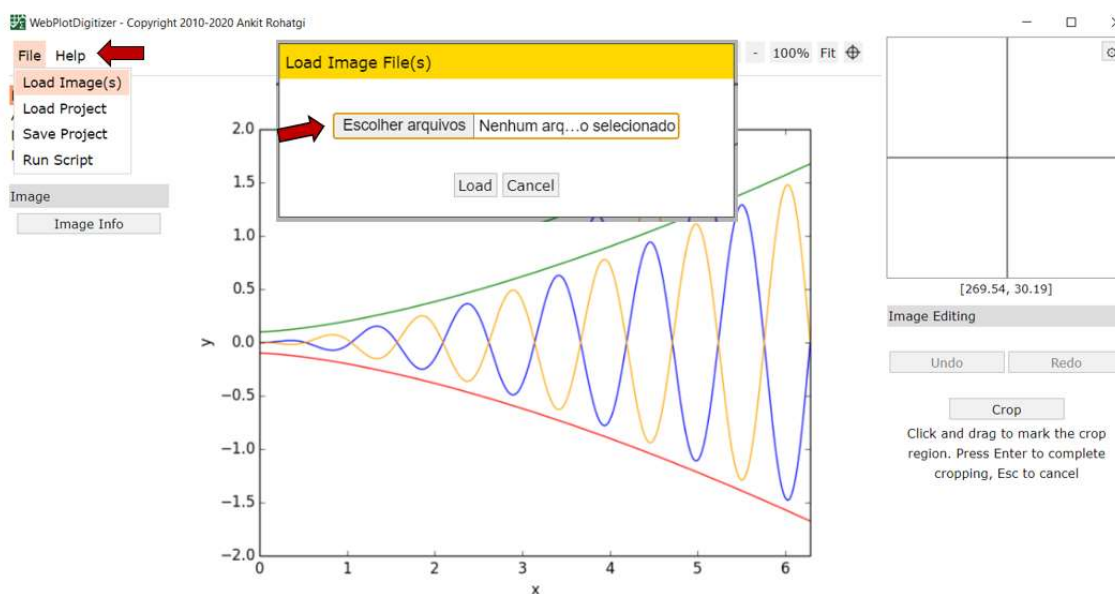
Copyright © 2010-2020, Ankit Rohatgi - Todos os direitos reservados.

Citação

Se desejar citar este software em seu trabalho, você pode usar as seguintes informações:

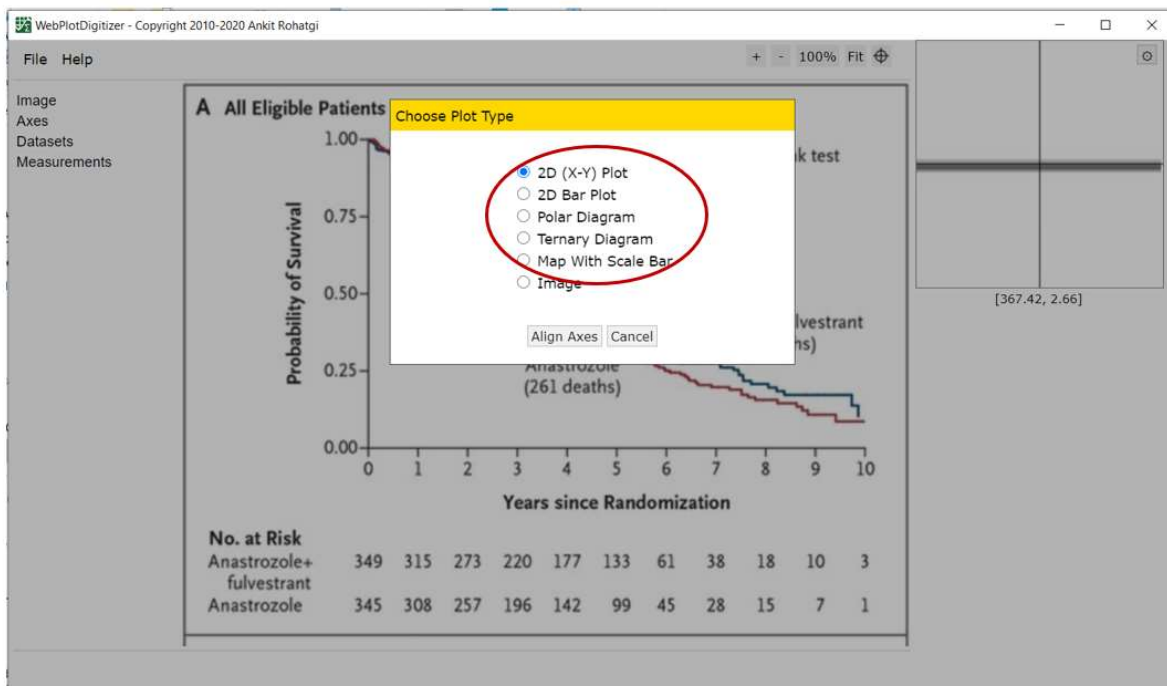
Autor: Ankit Rohatgi
 Título: WebPlotDigitizer
 Site: <https://automeris.io/WebPlotDigitizer>
 Versão: 4.4
 Data: novembro de 2020
 E-Mail: ankitrohatgi@hotmail.com
 Local: Pacífica, Califórnia, EUA

Passo 1: Inicie o *software* e abra a imagem do gráfico em “File > Load image(s) > Escolher arquivos”:

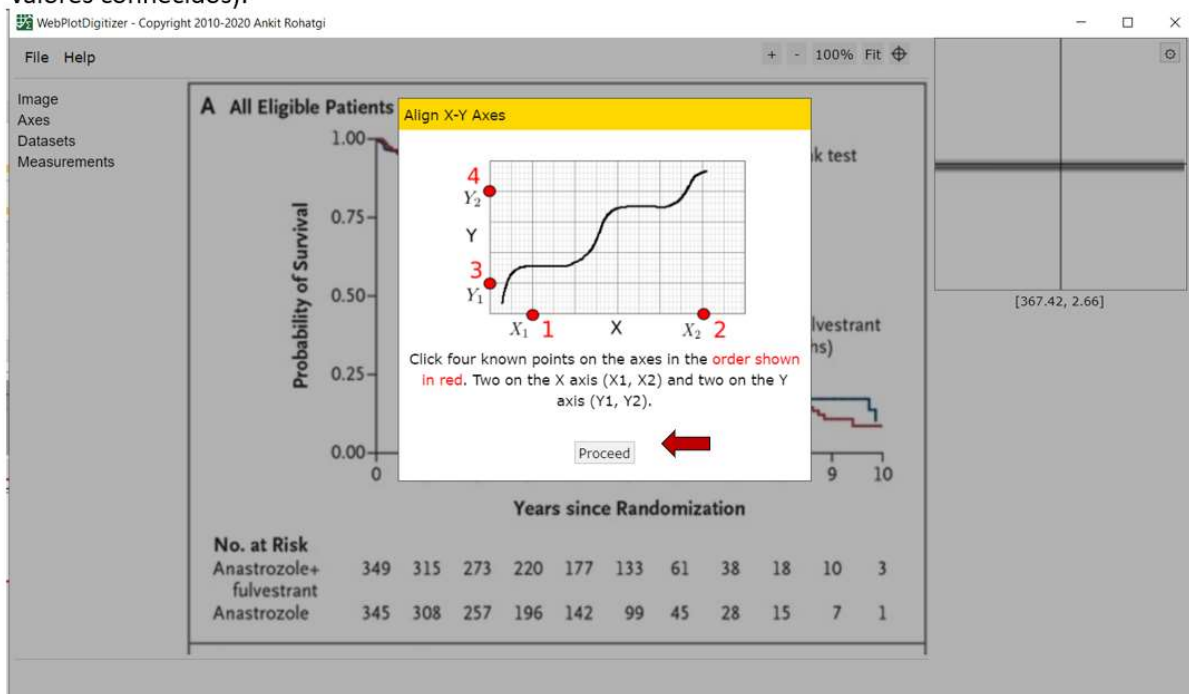


The screenshot shows the WebPlotDigitizer application window. The 'File' menu is open, with 'Load Image(s)' highlighted. A red arrow points to this menu item. The 'Load Image File(s)' dialog box is open over a plot, with 'Escolher arquivos' highlighted. Another red arrow points to this button. The plot shows a graph with x and y axes, a green trend line, and several oscillating lines in blue, yellow, and red. The 'Image Editing' panel on the right shows 'Undo' and 'Redo' buttons, and a 'Crop' button with instructions: 'Click and drag to mark the crop region. Press Enter to complete cropping, Esc to cancel'.

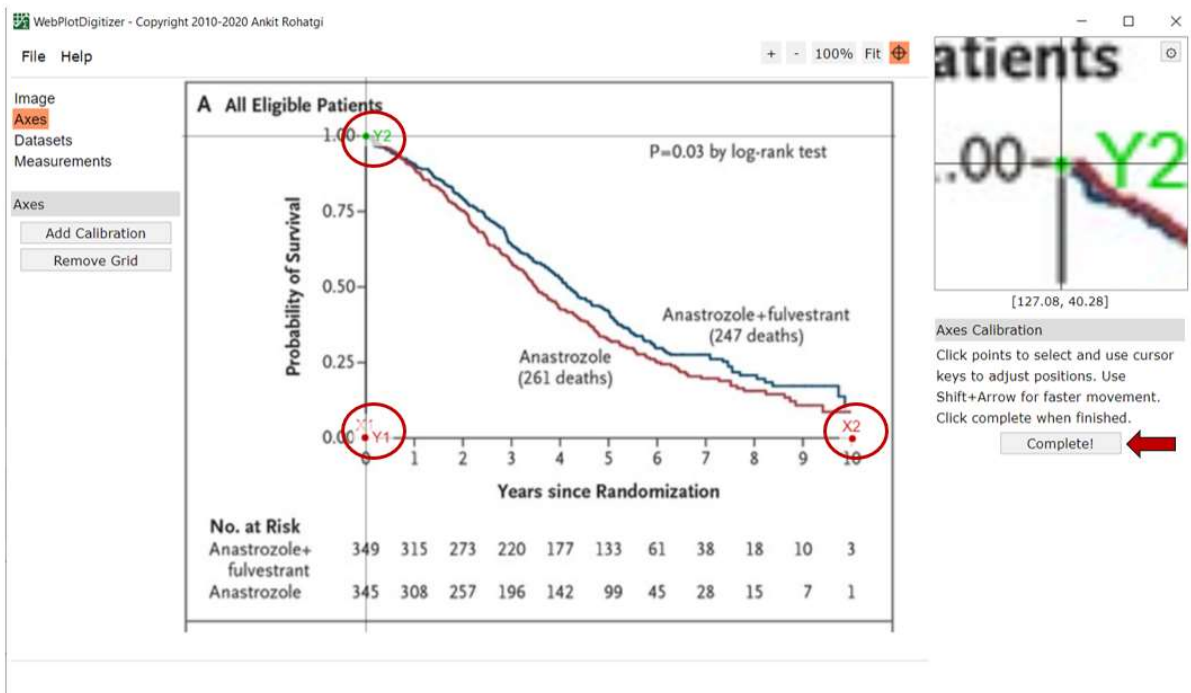
Passo 2: Selecione o tipo de gráfico (geralmente, a 1ª opção deve atender) e a opção “Align Axes” (alinhar eixos):



Passo 3: Observe as orientações de calibragem do programa para definir pontos conhecidos nos eixos e clique em “Proceed” (Obs: observe que não precisam ser coordenadas do mesmo ponto, apenas valores conhecidos):



Passo 4: Defina, primeiramente, dois pontos conhecidos do eixo X e, em seguida, dois do eixo Y. Por exemplo, neste gráfico, defini X_1 e X_2 como 0 e 10, respectivamente, e Y_1 e Y_2 como 0.00 e 1.00, respectivamente. Clique em “Complete!” no canto direito:



Passo 5: Preencha os valores dos pontos (Neste exemplo, os valores no “X-Axis” eram 0 e 10 e no “Y-Axis” eram 0.00 e 1.00) e clique em “OK”:

X and Y Axes Calibration

Enter X-values of the two points clicked on X-axis and Y-values of the two points clicked on Y-axis

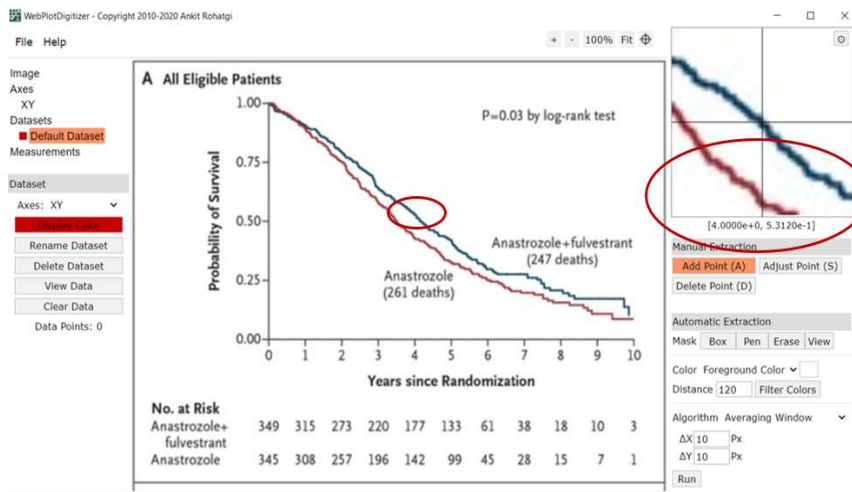
	Point 1	Point 2	Log Scale
X-Axis:	0	10	<input type="checkbox"/>
Y-Axis:	0.0	1.00	<input type="checkbox"/>

Assume axes are perfectly aligned with image coordinates (skip rotation correction)

*For dates, use yyyy/mm/dd hh:ii:ss format, where ii denotes minutes (e.g. 2013/10/23 or 2013/10 or 2013/10/23 10:15 or just 10:15). For exponents, enter values as 1e-3 for 10⁻³.

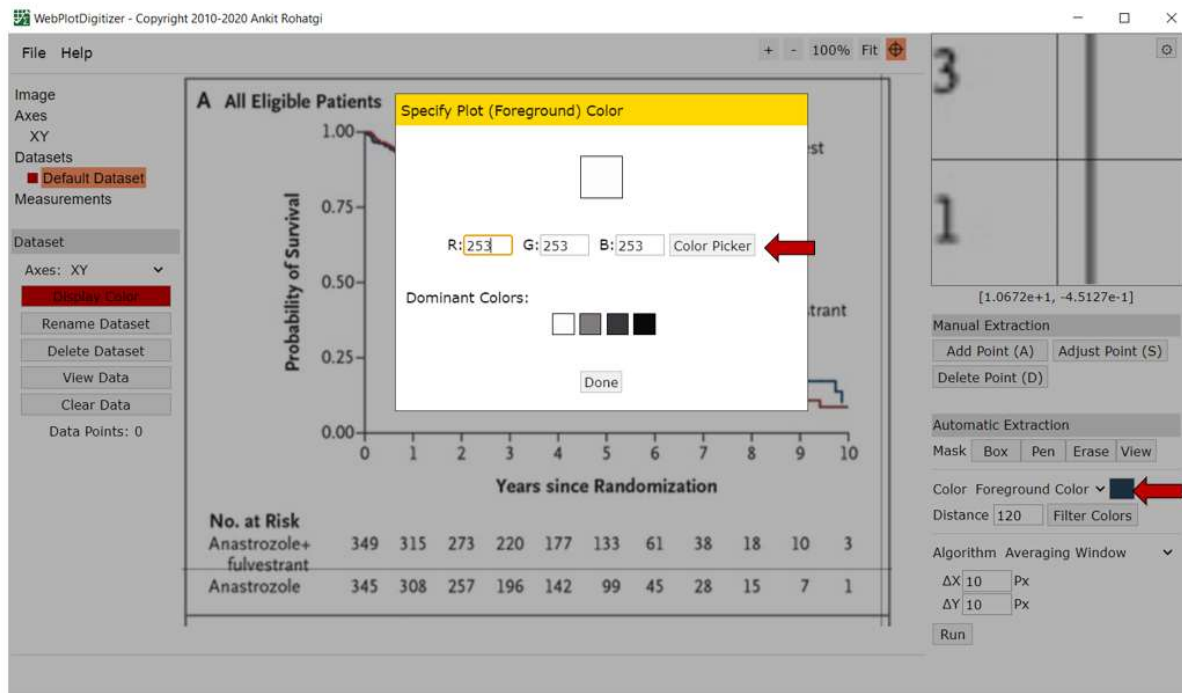
OK

Quase pronto! Seu gráfico está calibrado! Basta passar o mouse sobre qualquer ponto do gráfico que deseja conhecer e a coordenada (x, y) aparecerá no canto superior direito. Você pode solicitar uma sequência de vários pontos, seguindo mais alguns passos, nas telas seguintes.

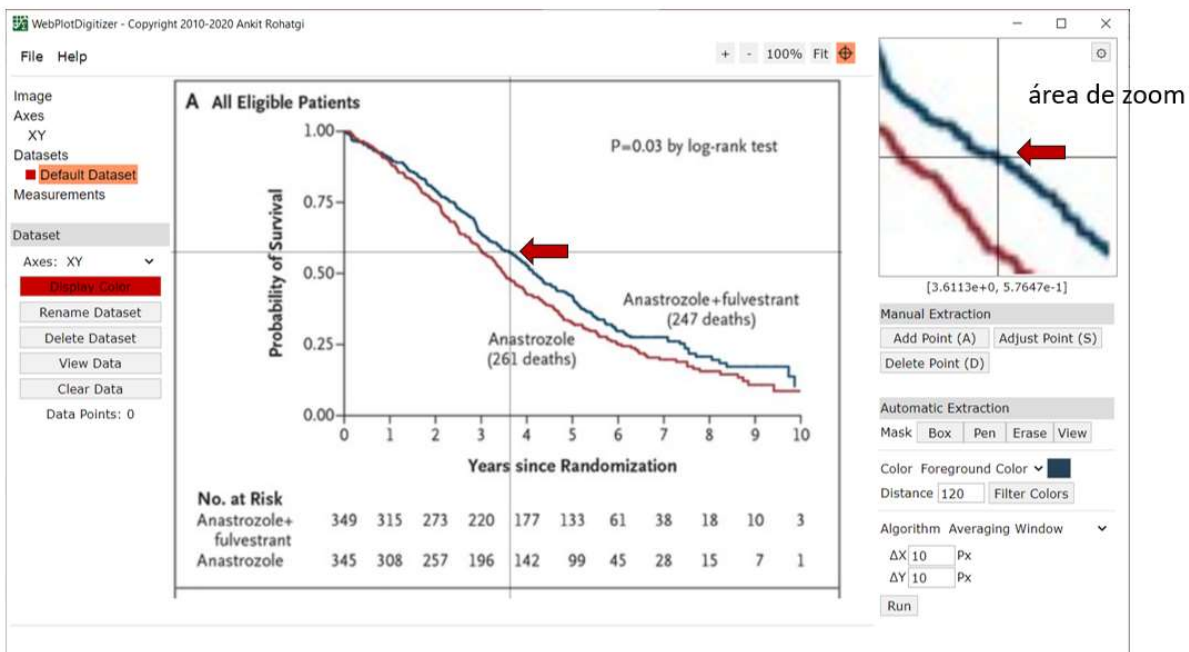


No exemplo ao lado, o programa informa que no ponto indicado, a coordenada X era 4 e a Y era 0,531, aproximadamente. Ou seja, no ano 4, a sobrevivência dos indivíduos era de 53% no grupo Anastrozol+Fulvestranto (dado que o estudo não fornecia).

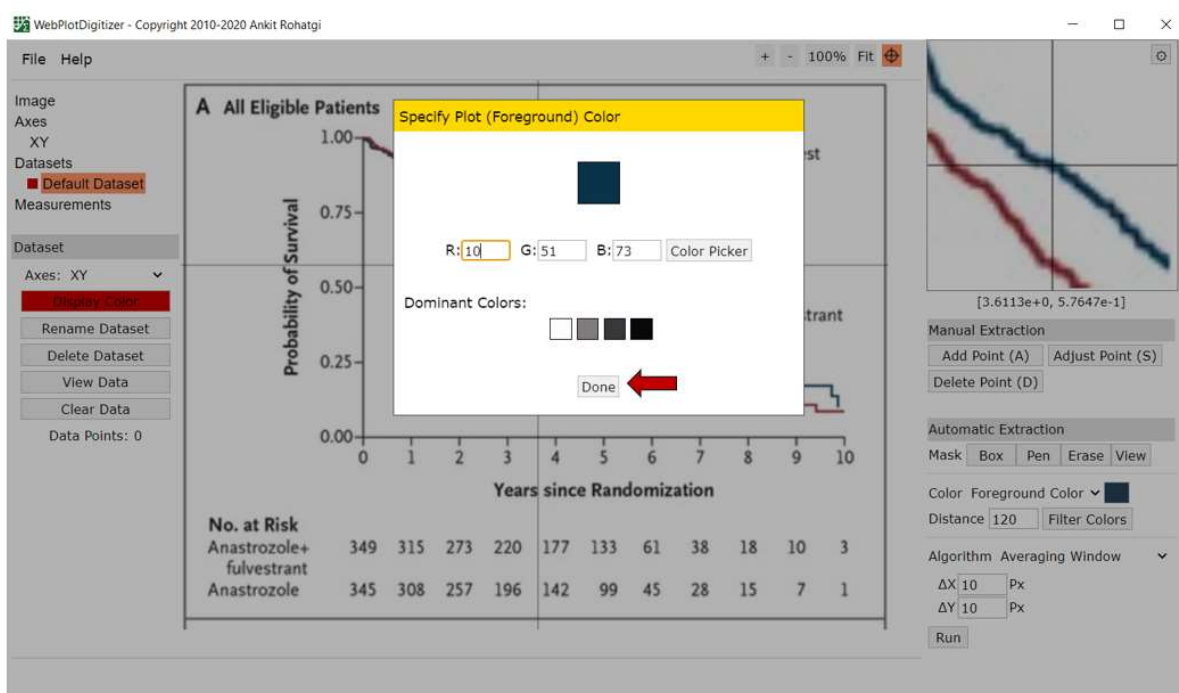
Passo 1: Na opção “Automatic Extraction”, especifique a cor da linha de dados, clique na caixa com cor, depois clique em “Color Picker”



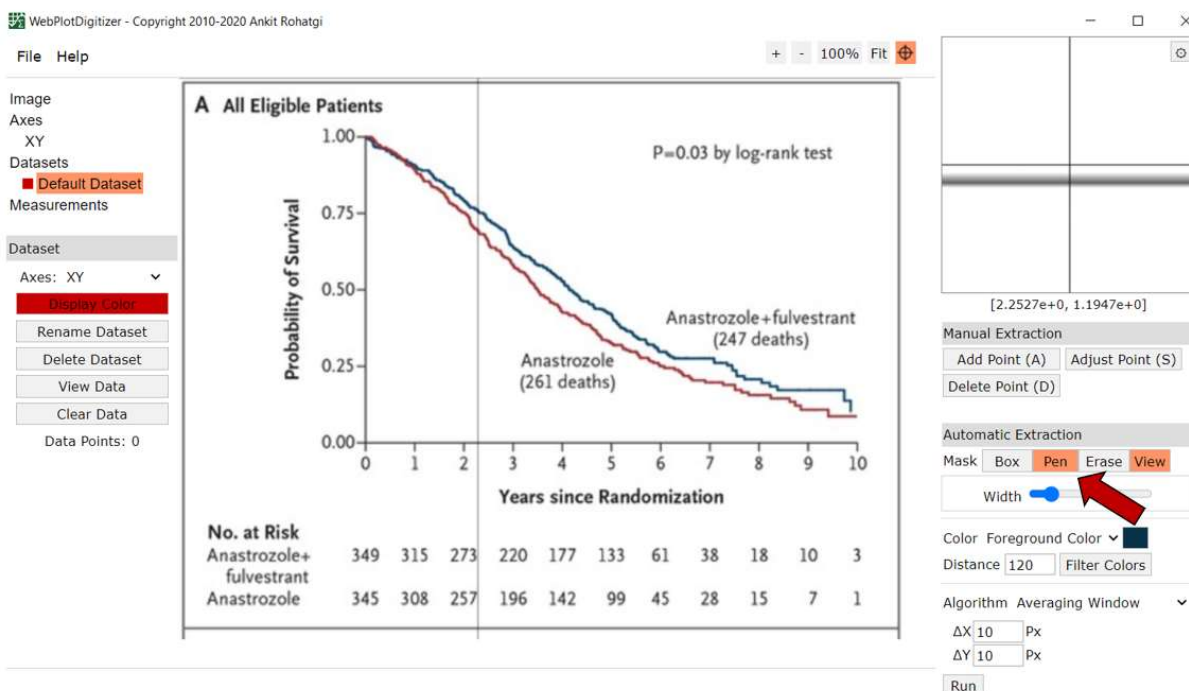
Passo 2: Passe o mouse na linha para capturar a cor, verifique sempre na área de zoom, se está em cima do ponto desejado



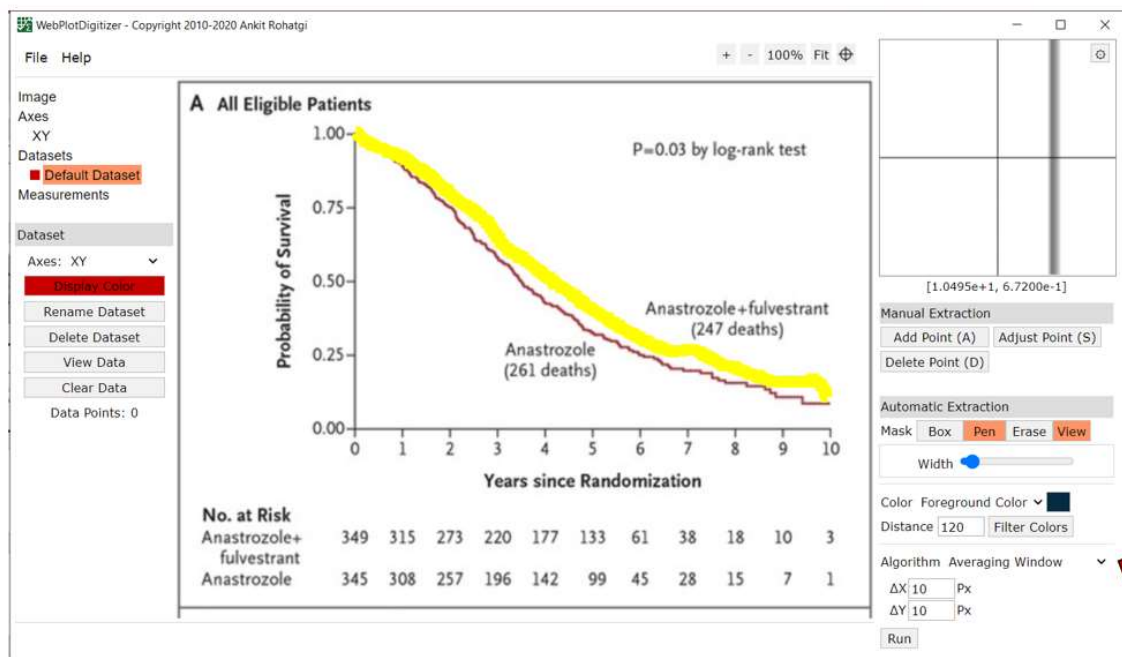
Passo 3: Após capturada a cor da linha o gráfico, clique em “Done”



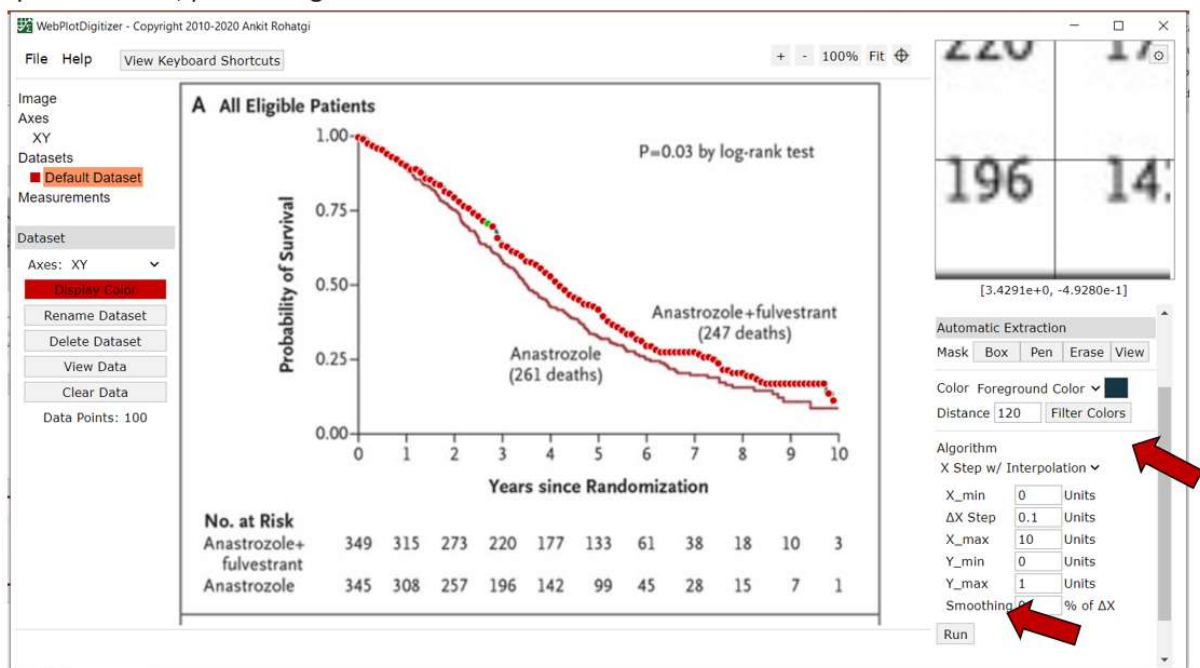
Passo 4: Após capturada a cor da linha o gráfico, escolha a opção “Pen” para extrair os pontos.



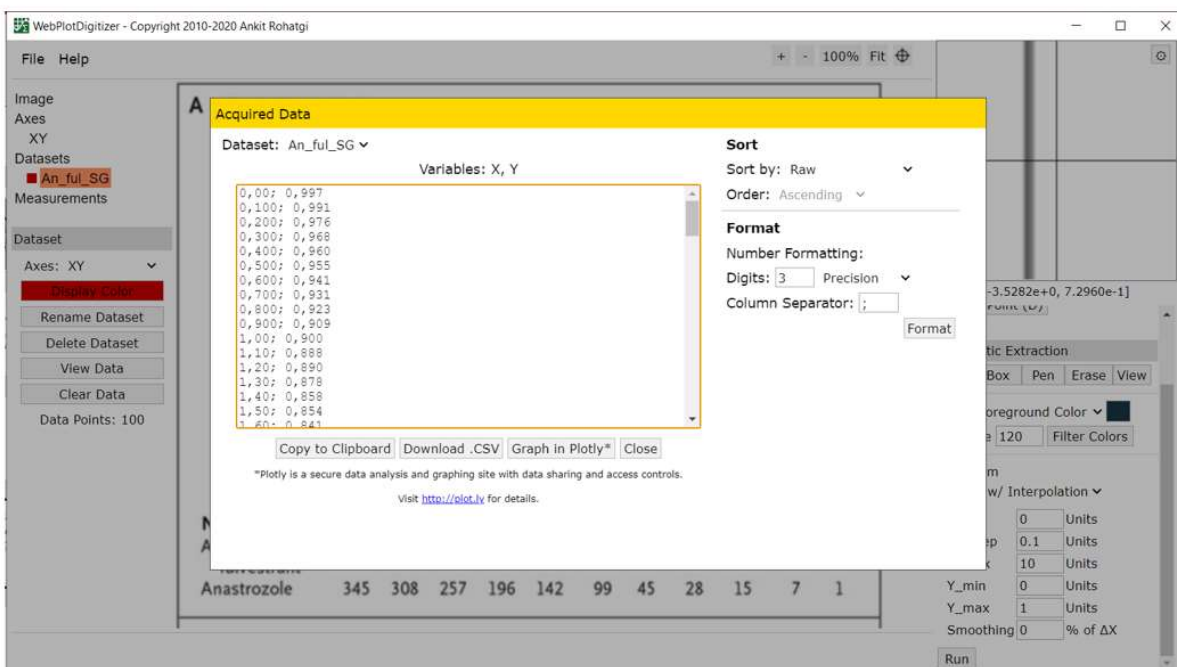
Passo 5: Depois de passar a caneta, escolha a forma automática de extração no “Algorithm”

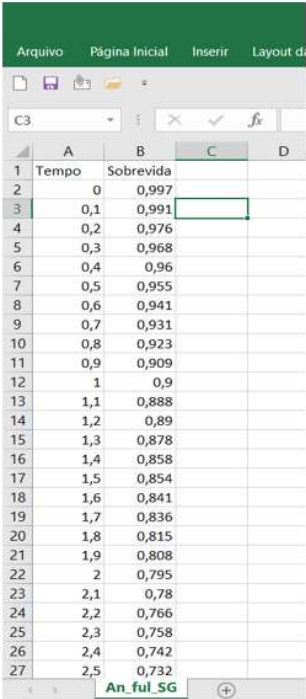


Passo 6: Escolha a forma automática de extração, e indique os intervalos de informação para o eixo x,y. Então clique em “Run”. Se algum ponto cair fora da linha, pode-se ajustar manualmente, clicando no “Adjust point(s)”. Utilize a área do zoom, para facilitar a visualização, dê um zoom na imagem sempre que necessário, para enxergar bem a curva.



Passo 7: Clicando em “View Data”, são fornecidas as coordenadas dos vários pontos (x, y). Salve em formato csv clicando em “Download .CSV” caso queira trabalhar com os dados no Excel ou outro software





The image shows a screenshot of the Microsoft Excel interface. The active window is titled 'An_ful_SG'. The spreadsheet contains two columns: 'Tempo' in column A and 'Sobrevida' in column B. The data is as follows:

	A	B	C	D
1	Tempo	Sobrevida		
2	0	0,997		
3	0,1	0,991		
4	0,2	0,976		
5	0,3	0,968		
6	0,4	0,96		
7	0,5	0,955		
8	0,6	0,941		
9	0,7	0,931		
10	0,8	0,923		
11	0,9	0,909		
12	1	0,9		
13	1,1	0,888		
14	1,2	0,89		
15	1,3	0,878		
16	1,4	0,858		
17	1,5	0,854		
18	1,6	0,841		
19	1,7	0,836		
20	1,8	0,815		
21	1,9	0,808		
22	2	0,795		
23	2,1	0,78		
24	2,2	0,766		
25	2,3	0,758		
26	2,4	0,742		
27	2,5	0,732		

Pronto! Salvando os dados em .CSV e trabalhando no Excel. Para trabalhar com o dado, basta abrir o arquivo no Excel e seguir o caminho no menu “Dados > Texto para Colunas > Delimitado > Vírgula > Concluir e fazer os ajustes de vírgulas e casa decimais. Agora os dados estão prontos para serem utilizados.

APÊNDICE F - Parâmetros das distribuições para os dados individualizados e agregados, para os dois tratamentos

Dados Individualizados

Distribuições	Anastrozol								
		Sobrevida global				Sobrevida livre de progressão			
Expo	rate	0.2083	0.1845	0.2352	0.0129	0.5770	0.5178	0.6429	0.0319
	shape	13.665	12.343	15.128	0.0709	0.9614	0.8851	10.443	0.0406
Weib	scale	47.484	43.450	51.894	0.2151	17.066	15.181	19.184	0.1019
	meanlog	12.026	10.904	13.147	0.0572	-0.00497	-0.12878	0.11884	0.06317
Lnormal	sdlog	10.111	0.9246	11.057	0.0461	116.728	108.023	126.135	0.04616
	mu	15.079	13.567	16.592	0.0772	0.15084	-0.04193	0.34361	0.09835
Ggen	sigma	0.7728	0.6570	0.9091	0.0640	113.245	104.069	123.230	0.04882
	Q	0.8426	0.4713	12.140	0.1895	0.28015	0.00671	0.55358	0.13951
Llog	shape	18.036	16.270	19.994	0.0948	15.109	13.807	16.534	0.0695
	scale	34.455	31.103	38.170	0.1800	10.335	0.9153	11.669	0.0640
Gomp	shape	0.1349	0.0775	0.1923	0.0293	-0.1322	-0.2034	-0.0609	0.0364
	rate	0.1442	0.1167	0.1780	0.0155	0.7115	0.6138	0.8247	0.0536
Distribuições	Anastrozol+Fulvestranto								
		Sobrevida global				Sobrevida livre de progressão			
Expo	rate	0.1786	0.1577	0.2023	0.0114	0.4486	0.4019	0.5007	0.0252
	shape	13414	12042	14942	0.0738	0.8826	0.8095	0.9623	0.0389
Weib	scale	54236	49410	59533	0.2579	21282	18722	24192	0.1391
	meanlog	13557	12252	14863	0.0666	0.1674	0.0309	0.3039	0.0697
Lnormal	sdlog	11588	10567	12708	0.0546	12848	11869	13908	0.0520
	mu	17728	16325	19131	0.0716	0.1941	-0.0537	0.4418	0.1264
Ggen	sigma	0.6528	0.5256	0.8108	0.0722	12810	11791	13917	0.0542
	Q	13175	0.8629	17721	0.2320	0.0439	-0.2971	0.3850	0.1740
Llog	shape	16859	15132	18783	0.0929	13315	12161	14577	0.0615
	scale	40375	36203	45028	0.2247	11995	10444	13776	0.0847
Gomp	shape	0.1561	0.0972	0.2149	0.0300	-0.1535	-0.2203	-0.0866	0.0341
	rate	0.1131	0.0900	0.1422	0.0132	0.6002	0.5134	0.7016	0.0478

Dados agregados

Distribuições	Anastrozol								
		Sobrevida global				Sobrevida livre de progressão			
Expo	rate	0.20848539	0.0010312164	20217424	0.000000e+00	0.61409012	0.002509358	244719999	0.000000e+00
	shape	130.420.095	0.0032429565	40.216.418	0.000000e+00	0.96219988	0.006102381	157.676.138	0.000000e+00
Weib	scale	0.13193616	0.0006808269	19.378.811	0.000000e+00	0.62902381	0.003757744	167.393.983	0.000000e+00
	mu_ln	0.63451852	0.0012283234	51.657.285	0.000000e+00	0.03030899	0.001693736	17894752	2,7469E-56
Lnormal	sd_ln	0.93041040	0.0016472271	56.483.432	0.000000e+00	102.271.211	0.001956868	522.627.082	0.000000e+00
	mu	145.387.309	0.0049520671	29.358.913	0.000000e+00	0.07182458	0.010624384	6760352	2,34587E-05
Ggen	sigma	0.81891420	0.0027301535	29995171	0.000000e+00	104.992.336	0.005527617	189.941.408	0.000000e+00
	Q	0.68786259	0.0148483216	4.632.595	2,8227E-230	0.02689636	0.025363182	1060449	289198,6
Llog	shape	191.199.418	0.0059799545	31.973.390	0.000000e+00	165.390.174	0.005726898	288.795.371	0.000000e+00
	scale	0.09545804	0.0008364615	11.412.125	0.000000e+00	0.88666024	0.004365262	203.117.298	0.000000e+00
Gomp	shape	0.15245892	0.0010363036	14711801	0.000000e+00	0.69578915	0.004121432	168822180	0.000000e+00
	rate	0.12025925	0.0022874409	5257371	1,9516E-265	-0.10868833	0.003798419	-28614097	8,1817E-126
Distribuições	Anastrozol+Fulvestranto								
		Estimate	Std..Error	t.value	Pr...t..	Estimate	Std..Error	t.value	Pr...t..
Expo	rate	0.17673822	0.0008718422	20271812	0.000000e+00	0.46477033	0.0024716642	18803943	0.000000e+00
	shape	131935070	0.0045076293	29269281	0.000000e+00	0.79330978	0.0032531894	24385601	0.000000e+00
Weib	scale	0.10447096	0.0008145886	12824996	0.000000e+00	0.56683320	0.0023626533	23991383	0.000000e+00
	mu_ln	0.72071975	0.0014887765	48410206	0.000000e+00	0.09973054	0.0008556145	11656013	0.000000e+00
Lnormal	sd_ln	0.92761330	0.0019588383	47355277	0.000000e+00	112444036	0.0009736030	115492691	0.000000e+00
	mu	162337829	0.0078108341	20783674	0.000000e+00	0.28975701	0.0061578174	4705515	2,0124E-245
Ggen	sigma	0.80708775	0.0045678555	17668854	0.000000e+00	128587528	0.0024268054	52986337	0.000000e+00
	Q	0.71447878	0.0239149894	2987577	9,9505E-134	0.16228166	0.0122338244	1326500	5,39617E-31
Llog	shape	192108754	0.0071939863	26704077	0.000000e+00	134232881	0.0029561080	45408652	0.000000e+00
	scale	0.06884839	0.0007960705	8648530	0.000000e+00	0.74425441	0.0024087143	30898410	0.000000e+00
Gomp	shape	0.12712189	0.0009591254	13253938	0.000000e+00	0.60253488	0.0011944818	50443202	0.000000e+00
	rate	0.11003414	0.0021959003	5010889	5,4896E-266	-0.16725602	0.0010599861	-15779076	0.000000e+00

ANEXO

1. Marguet S, Mazouni C, Ramaekers BLT, Dunant A, Kates R, Jacobs VR, et al. European cost-effectiveness study of uPA/PAI-1 biomarkers to guide adjuvant chemotherapy decisions in breast cancer. *Eur J Cancer* [Internet]. 2016;63:168–79. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S095980491632144X>
2. Vijayaraghavan A, Efrusy MB, Mazonson PD, Ebrahim O, Sanne IM, Santas CC. Cost-Effectiveness of Alternative Strategies for Initiating and Monitoring Highly Active Antiretroviral Therapy in the Developing World. *JAIDS J Acquir Immune Defic Syndr* [Internet]. 2007;PAP. Available from: <http://journals.lww.com/00126334-900000000-99696>
3. Walker S, Palmer S, Erhorn S, Brent S, Dyker A, Ferrie L, et al. Fludarabine phosphate for the first-line treatment of chronic lymphocytic leukaemia. *Health Technol Assess (Rockv)* [Internet]. 2009;13. Available from: <https://www.journalslibrary.nihr.ac.uk/hta/hta13Suppl106/>
4. Bederson JB, Awad IA, Wiebers DO, Piepgras D, Haley EC, Brott T, et al. Recommendations for the Management of Patients With Unruptured Intracranial Aneurysms. *Circulation* [Internet]. 2000;102:2300–8. Available from: <https://www.ahajournals.org/doi/10.1161/01.CIR.102.18.2300>
5. Harris L, Fritsche H, Mennel R, Norton L, Ravdin P, Taube S, et al. American Society of Clinical Oncology 2007 Update of Recommendations for the Use of Tumor Markers in Breast Cancer. *J Clin Oncol* [Internet]. 2007;25:5287–312. Available from: <http://ascopubs.org/doi/10.1200/JCO.2007.14.2364>
6. Pinilla-Dominguez P, Richardson J, Robertson J, Adam J. NICE guidance on bevacizumab in combination with gemcitabine and carboplatin for treating the first recurrence of platinum-sensitive advanced ovarian cancer. *Lancet Oncol* [Internet]. 2013;14:691–2. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S147020451370247X>
7. Mauskopf J, Chirila C, Masaquel C, Boye KS, Bowman L, Birt J, et al. RELATIONSHIP BETWEEN FINANCIAL IMPACT AND COVERAGE OF DRUGS IN AUSTRALIA. *Int J Technol Assess Health Care* [Internet]. 2013;29:92–100. Available from: https://www.cambridge.org/core/product/identifier/S0266462312000724/type/journal_article
8. Ghafoor Q, Sanghera P, Grieve RJ. A Study of Medical Intervention in Routine Breast Cancer Follow-up. *Clin Oncol* [Internet]. 2010;22:91–6. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0936655509003781>
9. Rich JS, Sox HC. Screening in the Elderly: Principles and Practice. *Hosp Pract* [Internet]. 2000;35:45–56. Available from: <http://www.tandfonline.com/doi/full/10.3810/hp.2000.10.215>
10. Mistry H, Nduka C, Connock M, Colquitt J, Mantopoulos T, Loveman E, et al. Venetoclax for Treating Chronic Lymphocytic Leukaemia: An Evidence Review Group Perspective of a NICE Single Technology Appraisal. *Pharmacoeconomics* [Internet]. 2018;36:399–406. Available from: <http://link.springer.com/10.1007/s40273-017-0599-9>

11. Farley C, Friedman D, Habtes I, Raskind-Hood C, Adams EK, Becker ER, et al. Screening Mammography in a Public Hospital Serving Predominantly African-American Women: A Stage–Survival–Cost Model. *Women’s Heal Issues* [Internet]. 2015;25:322–30. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1049386715000237>
12. Chew DP, Huynh LT, Liew D, Astley C, Soman A, Brieger D. Potential survival gains in the treatment of myocardial infarction. *Heart* [Internet]. 2009;95:1844–50. Available from: <http://heart.bmj.com/cgi/doi/10.1136/hrt.2009.174276>
13. Liu SL, Cheung WY. Role of surveillance imaging and endoscopy in colorectal cancer follow-up: Quality over quantity? *World J Gastroenterol* [Internet]. 2019;25:59–68. Available from: <https://www.wjgnet.com/1007-9327/full/v25/i1/59.htm>
14. Marshall A, Altman DG, Holder RL, Royston P. Combining estimates of interest in prognostic modelling studies after multiple imputation: current practice and guidelines. *BMC Med Res Methodol* [Internet]. 2009;9:57. Available from: <https://bmcmmedresmethodol.biomedcentral.com/articles/10.1186/1471-2288-9-57>
15. Smith TJ, Hillner BE. A Way Forward on the Medically Appropriate Use of White Cell Growth Factors. *J Clin Oncol* [Internet]. 2012;30:1584–7. Available from: <http://ascopubs.org/doi/10.1200/JCO.2011.39.9980>
16. Inadomi JM. Surveillance in Barrett’s Esophagus: A Failed Premise. *Keio J Med* [Internet]. 2009;58:12–8. Available from: <http://joi.jlc.jst.go.jp/JST.JSTAGE/kjm/58.12?from=CrossRef>
17. Mehta H. Coronary revascularization, when it needs to be done it needs to be done. *Eur Heart J* [Internet]. 2001;22:617–8. Available from: <https://academic.oup.com/eurheartj/article-lookup/doi/10.1053/euhj.2000.2494>
18. Jacobs I, Callanan V, Nichol G, Valenzuela T, Mason P, Jaffe AS, et al. The chain of survival. *Ann Emerg Med* [Internet]. 2001;37:S5–16. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0196064401655604>
19. Chung C, Christianson M. Predictive and prognostic biomarkers with therapeutic targets in breast, colorectal, and non-small cell lung cancers: A systemic review of current development, evidence, and recommendation. *J Oncol Pharm Pract* [Internet]. 2014;20:11–28. Available from: <http://journals.sagepub.com/doi/10.1177/1078155212474047>
20. Retèl VP, Joore MA, Knauer M, Linn SC, Hauptmann M, Harten WH van. Cost-effectiveness of the 70-gene signature versus St. Gallen guidelines and Adjuvant Online for early breast cancer. *Eur J Cancer* [Internet]. 2010;46:1382–91. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0959804910001620>
21. Hornberger J, Degtiar I, Gutierrez H, Shewade A, David Henner W, Becker S, et al. Cost-Effectiveness of Gene-Expression Profiling for Tumor-Site Origin. *Value Heal* [Internet]. 2013;16:46–56. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1098301512041307>
22. Klang SH, Hammerman A, Liebermann N, Efrat N, Doberne J, Hornberger J. Economic Implications of 21-Gene Breast Cancer Risk Assay from the Perspective of an Israeli-Managed Health-Care Organization. *Value Heal* [Internet]. 2010;13:381–7. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1098301510600708>

23. Oriana C, Martin H, Toby P, Chris C, Ruth G, Claudius R, et al. Complete Cytogenetic Response and Major Molecular Response as Surrogate Outcomes for Overall Survival in First-Line Treatment of Chronic Myelogenous Leukemia: A Case Study for Technology Appraisal on the Basis of Surrogate Outcomes Evidence. *Value Heal* [Internet]. 2013;16:1081–90. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1098301513019025>
24. Wang CJ, Cheng SH, Wu J-Y, Lin Y-P, Kao W-H, Lin C-L, et al. Association of a Bundled-Payment Program With Cost and Outcomes in Full-Cycle Breast Cancer Care. *JAMA Oncol* [Internet]. 2017;3:327. Available from: <http://oncology.jamanetwork.com/article.aspx?doi=10.1001/jamaoncol.2016.4549>
25. Parker C, Woods B, Eaton J, Ma E, Selby R, Benson E, et al. Brentuximab vedotin in relapsed/refractory Hodgkin lymphoma post-autologous stem cell transplant: a cost-effectiveness analysis in Scotland. *J Med Econ* [Internet]. 2017;20:8–18. Available from: <https://www.tandfonline.com/doi/full/10.1080/13696998.2016.1219358>
26. Doss S, Garrett Z, Sutcliffe F, Stevens A. NICE guidance on rituximab for first-line treatment of symptomatic stage III–IV follicular lymphoma in previously untreated patients. *Lancet Oncol* [Internet]. 2012;13:128–30. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1470204512700098>
27. Shafiq J, Delaney G, Barton MB. An evidence-based estimation of local control and survival benefit of radiotherapy for breast cancer. *Radiother Oncol* [Internet]. 2007;84:11–7. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0167814007000989>
28. Bast RC, Ravdin P, Hayes DF, Bates S, Fritsche H, Jessup JM, et al. 2000 Update of Recommendations for the Use of Tumor Markers in Breast and Colorectal Cancer: Clinical Practice Guidelines of the American Society of Clinical Oncology*. *J Clin Oncol* [Internet]. 2001;19:1865–78. Available from: <http://ascopubs.org/doi/10.1200/JCO.2001.19.6.1865>
29. Perrier L, Buja A, Mastrangelo G, Vecchiato A, Sandonà P, Ducimetière F, et al. Clinicians' adherence versus non adherence to practice guidelines in the management of patients with sarcoma: a cost-effectiveness assessment in two European regions. *BMC Health Serv Res* [Internet]. 2012;12:82. Available from: <https://bmchealthservres.biomedcentral.com/articles/10.1186/1472-6963-12-82>
30. Ridker PM. Clinical Application of C-Reactive Protein for Cardiovascular Disease Detection and Prevention. *Circulation* [Internet]. 2003;107:363–9. Available from: <https://www.ahajournals.org/doi/10.1161/01.CIR.0000053730.47739.3C>
31. Thomas DM. DES editorial. *Catheter Cardiovasc Interv* [Internet]. 2008;71:644–5. Available from: <http://doi.wiley.com/10.1002/ccd.21537>
32. Kleijnen S, Lipska I, Leonardo Alves T, Meijboom K, Elsada A, Vervölgyi V, et al. Relative effectiveness assessments of oncology medicines for pricing and reimbursement decisions in European countries. *Ann Oncol* [Internet]. 2016;27:1768–75. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0923753419358697>
33. Weintraub WS, Barnett P, Chen S, Hartigan P, Casperson P, O'Rourke R, et al.

Economics methods in the Clinical Outcomes Utilizing percutaneous coronary Revascularization and Aggressive Guideline-driven drug Evaluation (COURAGE) trial. *Am Heart J* [Internet]. 2006;151:1180–5. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0002870305008070>

34. Cosler LE, Hornberger J, Lyman GH. Economic analysis of targeting chemotherapy (CT) using a 21 gene RT-PCR assay in lymph node negative (LN-), estrogen receptor positive (ER+) early-stage breast cancer (ESBC). *J Clin Oncol*. 2004;22:6036–6036.

35. Monnickendam G, Zhu M, McKendrick J, Su Y. Measuring Survival Benefit in Health Technology Assessment in the Presence of Nonproportional Hazards. *Value Heal* [Internet]. Elsevier Inc; 2019;22:431–8. Available from: <https://doi.org/10.1016/j.jval.2019.01.005>

36. Pharmaceutical Benefits Advisory Committee. Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee. 2013; Available from: <http://www.pbac.pbs.gov.au/information/printable-version-of-guidelines.html>

37. Guerrero-c R, Parody G, Fe S, Osorio DI. Guía Metodológica Realización de Evaluaciones Económicas en el marco de Guías de Práctica Clínica - Versión final completa. 2014; Available from: [http://www.iets.org.co/Archivos/63/Guía Metodológica Realización de Evaluaciones Económicas en el marco de Guías de Práctica Clínica - Versión final completa.PDF](http://www.iets.org.co/Archivos/63/Guía%20Metodológica%20Realización%20de%20Evaluaciones%20Económicas%20en%20el%20marco%20de%20Guías%20de%20Práctica%20Clínica%20-%20Versión%20final%20completa.PDF)

38. (CSG) C de SG de M. 02 Guía para la conducción de estudios de evaluación económica. 2017; Available from: http://www.csg.gob.mx/contenidos/CB2013/cuadro_basico.html

39. Brasil. Ministério da Saúde. Secretaria de Ciência T e IED de C e T. Diretrizes metodológicas : Diretriz de Avaliação Econômica [Internet]. 2. ed. Brasília; 2014. Available from: http://conitec.gov.br/images/Artigos_Publicacoes/Diretrizes/AVE.pdf

40. Latimer N. NICE DSU technical support document 14: survival analysis for economic evaluations alongside clinical trials-extrapolation with patient-level data. Sheff Rep by Decis Support Unit [Internet]. 2011; Available from: [http://www.nicedsu.org.uk/NICE DSU TSD Survival analysis.updated March 2013.v2.pdf](http://www.nicedsu.org.uk/NICE%20DSU%20TSD%20Survival%20analysis.updated%20March%202013.v2.pdf)

41. CADTH. Canadian Agency for Drugs and Technologies in Health Agency [Internet]. 2006. Available from: https://www.cadth.ca/sites/default/files/pdf/186_EconomicGuidelines_e.pdf

42. Haute Autorité de Santé. Choices in Methods for Economic Evaluation: A Methodological Guide. Haute Aut Sante [Internet]. 2012;1–55. Available from: <https://www.has-sante.fr/upload/docs/application/pdf/2012->