MARIA DO ROSARIO DIAS DE OLIVEIRA LATORRE

**MORTALIDADE COMO PREDITOR DA
MORBIDADE**

DISSERTAÇÃO APRESENTADA À
FACULDADE DE SAÚDE PÚBLICA DA
UNIVERSIDADE DE SÃO PAULO
PARA OBTENÇÃO DO
TÍTULO DE MESTRE EM SAÚDE PÚBLICA
ORIENTADORA: PROF. DR. MARIA LÚCIA LEBRÃO

SÃO PAULO - 1992

À MINHA MÃE

AGRADECIMENTOS

À Lúcia, não só pela orientação efetiva, mas pela amizade e apoio
em diversos momentos difíceis no decorrer do mestrado.

Ao Jair, pela ajuda na elaboração do modelo matemático.

À Sabina e à Maria Helena que continuam a ser minhas super-mães.

Ao Fernão, que forneceu as listagens com os dados de mortalidade
e auxiliou na execução dos gráficos.

À Dayse, que fez a revisão nas referências bibliográficas.

E às diversas pessoas que me incentivaram durante toda a realização deste trabalho.

Meu muito obrigado a todos.

RESUMO

As doenças crônicas mostram-se como um desafio para o epidemiologista pela dificuldade de se conhecer a sua história natural, os agentes etiológicos, os suscetíveis e outras informações importantes para a sua prevenção, controle e tratamento. Dentre as doenças crônicas interessa, particularmente, o câncer devido à sua importância no obituário geral e pelos altos custos requeridos no seu tratamento.

Porém as estatísticas oficiais rotineiras dão pouca ou nenhuma informação sobre a sua incidência e/ou prevalência. Neste sentido, tendo em vista a parcialidade e a precariedade das estatísticas de morbidade, muitos estudos têm sido realizados com o objetivo de elaborar medidas indiretas de estimativas de incidência e prevalência.

O presente trabalho verifica a adequação de dois modelos matemáticos existentes na literatura que utilizam dados de população, mortalidade e sobrevivência para estimar a incidência de câncer de estômago, pulmão e mama feminina, utilizando dados do Município de São Paulo, para o ano de 1978. Além disso apresenta um modelo probabilístico que utiliza as mesmas informações dos modelos de literatura para estimar a incidência de doenças de caráter irreversível, com incidência, mortalidade e taxas de sobrevivência constantes por, pelo menos, um ano.

Os resultados mostraram que o modelo proposto por este trabalho apresentou as melhores estimativas dos coeficientes de incidência de câncer de estômago, pulmão e mama feminina, quando comparado com os modelos de literatura.

Os bons resultados obtidos incentivam a que se continue o aprimoramento de modelos matemáticos que utilizem dados de mortalidade para estimar a incidência de câncer.

ABSTRACT

Cronic diseases seem like a challenge for the epidemiologists, because it is difficult to know their natural history, etiologic agents, susceptibles and other important information for their prevention, control and treatment. Among cronic diseases, there is a special interest in cancer, because it is an important cause of death and the costs for its treatment are high.

However, the routine official statistics give little or no information at all about its incidence and/or prevalence. In relation to these lines, having in mind the partiality and the precariousness of the morbidity statistics, many studies have been carried out with the objective of elaborating indirect measures of estimations of incidence and prevalence.

This study checks the application of two mathematical models from the current literature that use population, mortality and survival data to estimate the incidence of stomach, lung and female breast cancer, using data from São Paulo City, for 1978. It was created a probabilistic model which uses the same informations of the literature models to forecast the incidence of the irreversible diseases, supposing that the rates of incidence, mortality and also survival are constants for at least one year.

The results have demonstrated that the proposed model of this study showed the best estimative for the cancer incidence rate for stomach, lung and female breast when compared to the literature models.

The good results obtained by this proposed model encourage the continuation of the development of mathematical models using mortality data to forecast cancer incidence rates.

ÍNDICE

	Pág.
I. Introdução.....	1
II. Modelos.....	16
III. Objetivos.....	29
IV. Material e Métodos.....	30
V. Resultados e Discussão.....	33
VI. Comentários.....	51
VII. Referências Bibliográficas.....	53
 Anexos.....	 57

1. INTRODUÇÃO

A preocupação com a saúde da população e a busca do completo bem estar físico, mental e social têm sido uma constante através dos tempos.

Aliada a esta inquietação, manifesta-se a necessidade de mensurar as condições de saúde de uma área, pois somente com o conhecimento de como as doenças se distribuem na população, quais os grupos de risco e sua variação no tempo e a inter-relação com as condições sócio-econômicas, culturais e ambientais da área pode-se planejar, adequadamente, medidas de prevenção e assistência à população, administrar os serviços e programas de saúde e distribuir recursos (12,31,50,56).

No entanto, a morbidade é um fenômeno muito difícil de ser estudado, pois é um fenômeno múltiplo. A definição de caso, isto é, o indivíduo doente, não é uniforme: não existe uma linha divisória clara entre saúde e doença (11,12,61) e ao invés de se dividir a população exposta ao risco de uma forma dicotômica doente-não doente, seria melhor analisar o processo mórbido por uma escala de valores entre os dois extremos (12,50).

Neste sentido, ao se pensar em realizar o diagnóstico das condições de saúde de uma comunidade, depara-se com um grande problema: como mensurar saúde?

As estatísticas oficiais rotineiras dão pouca ou nenhuma informação sobre incidência e/ou prevalência (64,65), na medida em que a maioria fornece informações sobre a demanda, acentuando o desconhecimento da população doente. Segundo LAURENTI e colaboradores (36), “embora as informações de morbidade provenham de várias fontes, todas elas pecam por serem incompletas ou inexatas. Desta forma, ainda que se somem todos os dados de morbidade de uma área, não se tem uma idéia completa e global de sua verdadeira morbidade”.

SWAROOP, em seu livro *Estadística Sanitaria* (61), ilustra muito bem a falha existente nos dados, pois não se tem indícios em relação a uma série de doenças e há duplicidade de informações de outras devido à sobreposição de diversas fontes de morbidade.

O fato de, quando se trabalha com morbidade, trabalhar-se com grandes massas de dados, tem acentuado o atraso nos estudos devido aos problemas com a coleta, armazenamento e manuseio dos dados e ao alto custo envolvido.

Algumas fontes de dados de morbidade, como as estatísticas hospitalares e as estatísticas de serviços de assistência médica, não têm base populacional e são uma função da demanda aos serviços de saúde. Essa demanda é determinada pelo uso que a população faz desses serviços, o que, de certa forma, mede a oferta dos mesmos (31) e não, obrigatoriamente, as

necessidades da população. Apenas a Notificação Compulsória e os Registros Especiais de Doenças fornecem a incidência de algumas patologias em determinada população, pois têm base populacional, mas referem-se a poucas doenças. Outras se referem a grupos populacionais específicos, como é o caso dos Registros Médicos de Indústrias. E outras, ainda, como os inquéritos populacionais, são muito caras.

Uma outra fonte de dados de morbidade muito utilizada é a Declaração de Óbito.

Os primeiros estudos elaborados objetivaram quantificar e qualificar os óbitos que ocorriam na Inglaterra em 1662, utilizando dados do Registro Geral (36). Porém, somente a partir de meados do século XIX é que as estatísticas vitais começaram a ser utilizadas de maneira mais adequada nas áreas de Saúde Pública e de Demografia, quando William Farr, chefe do Registro Geral da Inglaterra, mostrou as deficiências das estatísticas vitais e os métodos necessários para corrigi-las (36).

Desde então, quando se desejava fazer um diagnóstico das condições de saúde de uma comunidade, a ênfase sempre foi dada aos estudos de mortalidade. Isto se deve ao fato de que o óbito é um evento bem definido, tem caráter não renovável, o seu registro é obrigatório e, conseqüentemente, há sistemas de informação de mortalidade em quase todos os países, o que propicia a existência de longas séries históricas (5). Tais estudos podem ser feitos segundo importantes variáveis epidemiológicas que constam da Declaração de Óbito, como, por exemplo, sexo, idade, local de ocorrência, causa básica do óbito e outras (36,54,57,58,59).

Devido à existência do **Sub-Sistema de Informações de Mortalidade** implantado no Ministério da Saúde, há uma série de estudos nesta área que procuram analisar tendências ou, simplesmente, descrever séries históricas. No entanto, os estudos de mortalidade referem-se apenas à causa básica do óbito, isto é, “a doença ou lesão que iniciou a cadeia de acontecimentos patológicos que conduziram diretamente à morte, ou as circunstâncias do acidente ou violência que produziram a lesão fatal” (36); o que os torna inadequados quando se pensa do ponto de vista de morbidade. Isso ocorre não só pelo problema existente em muitas áreas geográficas, devido a uma alta porcentagem de óbitos por causas mal definidas mas, também, pelo fato de que muitas doenças não levam o indivíduo diretamente ao óbito fazendo com que fiquem sub-enumeradas (56).

Há uma diminuição dessa imprecisão quando os estudos são feitos utilizando-se as causas múltiplas de morte. Quando se estuda o óbito através deste sistema, pode-se analisá-lo como resultado de muitas causas simultâneas ou seqüenciais (36,54,57,58,59), visto que estas causas associadas (ou desencadeantes) também podem ser mensuradas. No Estado de São Paulo, estas informações estão disponíveis na Fundação SEADE a partir de 1983, quando foi introduzido o Programa de Classificação Automática de Causas de Morte, adaptado do similar norte-americano ACME (*Automated Classification of Medical Entities*). Este programa seleciona, automaticamente, a causa básica do óbito, de acordo com as regras internacionais de seleção e modificação de causa

básica e armazena, para cada óbito, todas as causas mencionadas no atestado de óbito, até um máximo de sete causas (57,58,59). Isso possibilita que a mortalidade seja estudada através das causas múltiplas de morte, o que maximiza a utilização do dado, fazendo com que o atestado de óbito passe a ser um instrumento mais útil nos estudos de morbidade. As tabulações por causas múltiplas podem ser feitas pela frequência total das causas mencionadas nos atestados (seja como causa básica, seja como causa associada) ou, então, por associação de causas (que mostra, por exemplo, em relação a uma determinada causa básica, quais as causas consequenciais e/ou contributivas mais frequentes e vice-versa). Os estudos feitos utilizando o ACME têm sido, principalmente, em relação à mortalidade infantil e à prevalência de diabetes mellitus e hipertensão em adultos (54).

No entanto, o óbito é um evento final de um processo dinâmico mais amplo (5). Em que pese o grande valor dos estudos de mortalidade, eles não refletem, totalmente, a incidência e a prevalência das diversas doenças em uma população (6,56).

Neste sentido, tendo em vista a parcialidade e a precariedade das estatísticas de morbidade, muitos estudos têm sido realizados com o objetivo de elaborar medidas indiretas de estimativas da incidência e da prevalência de algumas doenças (15).

Desde o início do século tem havido a preocupação em tentar “modelar” as informações existentes a respeito da mortalidade e sobrevivência de algumas doenças. O objetivo principal era descobrir uma função matemática ou uma distribuição de probabilidades conhecida que pudesse explicar os dados existentes. Fazia-se a comparação entre as curvas de mortalidade com a função exponencial, ou logarítmica ou, então, com a distribuição de Poisson ou outra para verificar qual a curva que melhor se adaptava aos dados (30). Um exemplo deste tipo de estudo foi a relação linear observada entre o logaritmo da incidência de diversos tipos de câncer e o logaritmo da idade do óbito (3,4).

Porém, isso era extremamente complicado, visto que quanto mais parâmetros eram utilizados, menos operacional ficava o modelo. Somente com o surgimento do computador é que a metodologia quantitativa pode ser melhor utilizada e foi possível dar maior liberdade às “fantasias” matemáticas (40).

Os objetivos principais de um modelo são estudar o comportamento de uma população sob o ponto de vista de determinada doença, prever o comportamento da mesma, simular diversas situações futuras não necessariamente vistas e/ou medidas e descobrir relações não aparentes. Para tanto o modelo deve procurar retratar, o melhor possível, a realidade. E é aí que aparece o primeiro problema quando se trabalha com modelagem: quanto mais real é o modelo, do ponto de vista biológico, mais difícil é o seu tratamento matemático (25,31,40,51), tornando-se, às vezes, tão complexo que perde a sua utilidade. Por isso devem ser feitas simplificações dessa realidade. A esse respeito, NEYMAN (apud Meyer, 42), sabiamente, escreveu: “o bom resultado do modelo depende de que os pormenores desprezados sejam realmente sem importância na

elucidação do fenômeno estudado”.

Neste ponto, geralmente, há um dilema entre o matemático (que tenta colocar o máximo de axiomas e postulados) e o epidemiologista (que deseja simplificar o máximo possível o conjunto de processos lógicos). Segundo o matemático NOONEY (51) “somente não é possível o diálogo entre o matemático e o anti-matemático” e sempre será possível apresentar um modelo que descreva a história natural da doença. Este é um processo contínuo de abstração, predição, verificação e refinamento do modelo.

Os modelos matemáticos são os que melhor se apresentam para representar a história natural da doença, pelo fato de serem estruturais e não funcionais, como os modelos estatísticos. Os modelos estatísticos são restritivos devido às diversas suposições que devem ser satisfeitas e, muitas vezes, algumas de suas parcelas ou fatores não têm explicação biológica ou epidemiológica (como, por exemplo, o produto de termos em um modelo multivariado) (29). Os modelos estatísticos são mais indicados quando utilizados para estudar medidas de efeito. Porém, nesse caso, não é possível estudar a ação biológica desses fatores e nem o mecanismo biológico da doença. No que se refere aos modelos preditivos de séries temporais, eles podem ser empregados quando se deseja prever comportamentos futuros de séries de mortalidade ou morbidade já existentes (35).

Vários trabalhos têm sido publicados utilizando modelos estatísticos para estimar a incidência de uma doença em uma população, porém todos sofrem restrições do ponto de vista conceitual.

Em 1978, COCHRANE e colaboradores (17) desenvolveram um modelo de regressão múltipla, cujas hipóteses poderiam ser aventadas quando se considera um modelo para estimar morbidade, embora o objetivo daquele estudo fosse estimar alguns coeficientes de mortalidade. Utilizaram como variáveis independentes fatores relacionados aos serviços de saúde, consumo alimentar e a indicadores econômicos e demográficos. Eles se basearam no fato de os países desenvolvidos apresentarem diferenças quanto aos seus coeficientes de mortalidade e aos seus gastos *per capita* com a saúde, apesar de o pessoal médico e paramédico terem graus homogêneos de conhecimentos e todas as pessoas da população terem à sua disposição praticamente os mesmos recursos. Selecionaram, segundo alguns critérios, 18 países considerados desenvolvidos para fazerem parte do estudo. As variáveis independentes foram selecionadas pela disponibilidade de dados e pela opinião, subjetiva, de que estivessem relacionadas com o estado de saúde da comunidade. A análise foi feita considerando como variáveis dependentes os coeficientes específicos por idade até a idade de 64 anos, os coeficientes de mortalidade materna, mortalidade perinatal e o de mortalidade infantil. As análises foram feitas para dois anos distintos, 1960 e 1970, obtendo-se os mesmos resultados: os modelos foram bons para estimar alguns coeficientes como, por exemplo, a mortalidade materna e a mortalidade infantil. Porém deram alguns resultados inesperados como, por exemplo, forte correlação positiva entre a prevalência de pediatras e a

mortalidade para as idades mais jovens, inclusive a infantil. O estudo foi prejudicado pelo fato de que os 18 países escolhidos intencionalmente não eram homogêneos, tanto interna quanto externamente, em relação aos fatores estudados.

Muitas críticas foram feitas ao trabalho desenvolvido por COCHRANE e colaboradores. GRAVELLE & BACKHOUSE, em 1987, (28) fizeram uma análise da metodologia utilizada, verificando entre outras coisas, o possível viés existente devido ao fato de que outras variáveis que não foram contempladas no modelo poderiam estar influenciando na utilização dos serviços de saúde, ou então, de que as variáveis escolhidas poderiam estar agindo de forma diferente em cada país. Outro fato analisado foi a multicolinearidade das variáveis independentes, pois algumas delas poderiam estar sendo influenciadas por uma outra variável do modelo. Em relação à definição, seleção e omissão das variáveis explicativas, os critérios de inclusão e exclusão deveriam ser mais objetivos. Lembraram, também, o fato de que as variáveis poderiam não ter permanecido constantes nos países nos dois anos considerados. Replicaram o estudo para o ano de 1977 e concluíram que o modelo é pouco seguro e necessitaria de maior elaboração para que os resultados pudessem ser interpretados adequadamente.

DAMIANI e colaboradores, em 1982, (21) propuseram um modelo de regressão linear múltipla para estimar a morbidade de algumas doenças (variável dependente) a partir de dados de mortalidade e de população (variáveis independentes), partindo do pressuposto de que os doentes de uma dada afecção correspondem à parte da população estatisticamente exposta ao risco de adquirir a afecção. Os coeficientes de incidência e de prevalência, segundo sexo e idade para a França, foram estimados para as seguintes afecções: tuberculose (todas as formas), câncer (algumas localizações), doenças cérebro-vasculares, doenças do coração, bronquite, pneumonia, diabetes, alcoolismo, cirrose do fígado, nefrite e nefrose, suicídios e acidentes. Porém, nem todos os resultados puderam ser testados devido à inexistência de dados para a França. Apenas os resultados de câncer foram comparados adequadamente, mostrando que os valores obtidos pelo modelo são quatro vezes maiores do que os registrados. Este modelo faz algumas suposições que não correspondem ao que se conhece na literatura, como, por exemplo, que o coeficiente anual de incidência pela doença é igual ao coeficiente de mortalidade por essa causa (isto é, todos os indivíduos que adquiriram a doença irão morrer no mesmo ano) ou, então, que o relacionamento entre a morbidade e mortalidade é linear com a idade.

Em outro estudo foi proposta uma outra alternativa para se descrever a morbidade de uma área pelo *The Resource Allocation Working Party (RAWP) do Department of Health and Social Security, Birmingham, England* (1976), que produziu um relatório que aconselhava o uso do *Standardised Mortality Ratio (SMR)* como indicador de morbidade quando se quisesse fazer estudos para planejamento de alocação de recursos na área da saúde, considerando que os SMR's de cada regional de saúde são indicadores das necessidades de recursos de saúde da população da área considerada (8,26,28). Utilizar o SMR como indicador de morbidade é muito criticável, pois a mortalidade, conforme já comentado anteriormente, não descreve, de maneira adequada, a

frequência real das doenças na área considerada.

Ainda analisando o SMR, FORSTER, em 1977, (26) testou a adequação da mortalidade, medida através dos SMR's regionais, como indicador das necessidades de uma área, comparando-a com quatro indicadores de morbidade obtidos através do *General Household Survey*. A análise foi feita considerando a ordem de importância dos quatro indicadores de morbidade e dos SMR's, através do teste de postos de Spearman. Houve diferença significativa nas necessidades das regionais medidas pela mortalidade ou pela morbidade, concluindo-se que o SMR não é um bom indicador de morbidade. Apenas foi encontrada correlação significativa entre a mortalidade e a morbidade para as doenças crônicas.

Devido aos baixos coeficientes de mortalidade apresentados pelas doenças infecciosas, GANDY, em 1979, (27) propôs como medida alternativa, quando se quisesse fazer estudos de alocação de recursos em áreas onde há grande número de doenças infecciosas, o *Standardised Notification Rate* (SNR). Para demonstrar a sua hipótese, escolheu a tuberculose pulmonar por ser uma doença infecciosa com grande número de óbitos quando comparada com outras doenças de notificação compulsória, na Inglaterra. O SNR, no caso da tuberculose pulmonar, mostrou ser uma medida mais adequada do que o SMR para se estabelecer prioridades. A restrição feita ao SNR é que ele se baseia na notificação compulsória de doenças e depende, portanto, da qualidade da notificação do local em que é feita, além de não contemplar todas as doenças infecciosas.

Pode-se observar que estes exemplos citados mais parecem exercícios estatísticos do que propriamente uma tentativa de utilização de métodos indiretos para estimar a morbidade, além de estarem sujeitos ao risco de incorrer em uma falácia epidemiológica, o que mostra a vantagem de se utilizar modelos matemáticos quando se deseja descrever o processo mórbido por meio de modelos .

Os modelos matemáticos são de dois tipos: o determinístico e o probabilístico (ou estocástico) (40,42).

De uma maneira resumida, diz-se que um modelo é determinístico quando ele prevê o resultado em termos de números absolutos. Já os modelos probabilísticos descrevem a função distribuição de probabilidades para os possíveis resultados da variável aleatória. Exemplificando: um modelo determinístico fornece o número de casos novos de uma doença e o estocástico, a probabilidade de surgirem casos novos (40,42). Pode-se dizer que um modelo probabilístico é mais realista quando se trabalha com morbidade, pois a saúde de um indivíduo está sujeita a riscos que são regidos por leis de probabilidade. Porém os modelos que mais contribuíram para a compreensão das doenças foram os determinísticos por terem um tratamento matemático mais fácil. O primeiro impulso dado aos estudos utilizando modelos estocásticos se deu após a década de 40, quando a teoria da probabilidade teve grande desenvolvimento. O segundo, a partir da década de 70, com a difusão de uma ferramenta indispensável: o computador.

Ao se definir que modelo matemático será empregado, deve-se analisar, primeiramente, se o fenômeno a ser estudado é de natureza determinística ou probabilística. Porém esta análise pode servir apenas como orientação, devido ao fato de que algumas situações claramente probabilísticas podem ser descritas, de maneira satisfatória, por um modelo determinístico e vice-versa.

Além disso, podem ser construídos modelos matemáticos que incorporem os dois elementos (determinísticos e probabilísticos) sempre que os conceitos teóricos não se conflitarem e a imaginação conseguir elaborar.

As abordagens para doenças transmissíveis e não transmissíveis são feitas de formas diferentes, devido ao maior conhecimento que se tem a respeito da história natural das primeiras. Ao se construir um modelo para uma doença transmissível ele vai depender da doença, dos mecanismos de transmissão, do comportamento do agente, do número de suscetíveis, do número de infectados e outros pontos que não precisam ser considerados quando se trata de uma doença não transmissível (24,40). Como ainda existem pontos obscuros sobre o início do processo mórbido e o seu desenrolar quando se fala em doenças não transmissíveis, a abordagem é mais genérica.

DAMIANI, em 1973, (19) propôs um modelo estocástico único para estimar morbidade tanto para doenças transmissíveis quanto para não transmissíveis a partir das estatísticas de mortalidade na França. O modelo calcula a incidência da doença por idade e a prevalência no ponto, também por idade, para três doenças: tuberculose pulmonar, diabetes mellitus e alcoolismo (incluindo cirrose do fígado), utilizando série histórica de óbitos, segundo idade e sexo. Este é um modelo puramente teórico e as hipóteses para estimar os parâmetros de transição foram feitas sem justificativas biológicas, pois não se aplicam às doenças transmissíveis e não transmissíveis ao mesmo tempo. Foram utilizados dados do Seguro Social da França e dados de literatura para verificar a adequação do modelo (19,20) e as estimativas obtidas só foram boas para diabetes mellitus, embora as curvas de incidência das patologias estudadas, segundo sexo e idade, tivessem formatos comparáveis aos encontrados na literatura.

Com a transição epidemiológica ocorrida neste século, através do controle das doenças transmissíveis, observou-se um aumento da esperança de vida em todos os países. Este fato faz com que as doenças crônico-degenerativas passem a ter uma importância maior quando se pensa em morbidade.

As doenças crônicas mostram-se como um desafio para o epidemiologista pela dificuldade de se conhecer sua história natural, os agentes etiológicos, os suscetíveis e outras informações importantes para a sua prevenção, controle e tratamento. Dentre as doenças crônicas interessa, particularmente, o câncer devido à sua importância no obituário geral (no Brasil é a terceira causa de morte, após as doenças do aparelho circulatório e das causas externas (43)) e pelos

altos custos requeridos no seu tratamento.

BOSCHI-PINTO, 1991, descreve, detalhadamente, o aumento do peso dos neoplasmas malignos no perfil da morbi-mortalidade no Brasil, em sua monografia sobre mortalidade por câncer (6), o que reforça a importância das neoplasias malignas quando se pensa em conhecer as doenças de maior incidência no Brasil.

Os dados de incidência de câncer são obtidos, de forma completa, através dos mais de 200 Registros de Câncer dispersos pelo mundo (63). Nos Estados Unidos, são publicadas, anualmente, estatísticas a respeito da incidência, sobrevivência e mortalidade de câncer pelo Instituto Nacional de Câncer, coletados por 10 Registros de Câncer de áreas geográficas selecionadas que representam 10% da população americana (14). Nos países que fazem parte do Mercado Comum Europeu há cerca de 70 registros populacionais de câncer que cobrem cerca de um terço da população (aproximadamente 310 milhões de pessoas) (63). Isso faz com que a preocupação maior dos países mais desenvolvidos seja em relação à precisão da informação (14,24,37,54).

Em que pesem os problemas inerentes a um Registro Especial de Doenças (financiamento, cobertura, fidedignidade dos dados e outros), provavelmente existem no mundo mais informações sobre o câncer do que para qualquer outra doença não transmissível e, em alguns países como Inglaterra e País de Gales (18,60), Estados Unidos (7), Suíça (65), Noruega (33,52), Finlândia (32,33) e outros do continente europeu, além do Canadá, Austrália (63) e Japão há diversos bons Registros de Câncer há muito tempo que têm auxiliado, entre outras coisas, na determinação da incidência e no estudo da sobrevida de pacientes portadores de câncer, segundo sexo, idade e estágio da doença, o que permite a definição de metas e programas de prevenção e controle com maior precisão (7).

O mesmo já não ocorre no Brasil, particularmente no Município de São Paulo, onde apesar de as neoplasias malignas serem responsáveis por 13% dos óbitos (2), o Registro de Câncer de São Paulo, existente desde 1969, está sempre sofrendo solução de continuidade devido a problemas financeiros e os últimos dados publicados referem-se ao ano de 1978 .

Mesmo com fartas informações sobre as neoplasias, os modelos matemáticos iniciais para estimar a incidência de câncer foram desenvolvidos por estudiosos americanos e ingleses e trataram o câncer como um processo dinâmico e complexo de crescimento e diferenciação de células (1,9,10,22,37,46,48,49,62) .

O primeiro modelo matemático para carcinogênese foi construído por ARMITAGE & DOLL (1954) (3). Este é um modelo estocástico e estabelece a hipótese de que o início do processo de carcinogênese ocorre quando uma célula-mãe normal se transforma em maligna devido a um componente hereditário . Além disso supõe que o logaritmo dos coeficientes de

incidência específicos por idade para alguns tipos de câncer aumentam linearmente com o logaritmo da idade. Em 1957 (9,46,48,62), ampliaram este modelo considerando que o início do desenvolvimento do câncer se dava em dois estágios : o primeiro onde a célula-mãe normal se transforma em células intermediárias através de um crescimento exponencial e o segundo estágio onde este conjunto de células pré-neoplásicas se transforma, sem controle, em células malignas até atingir um determinado volume detectável e se transformar em tumor. A transformação é instantânea, porém o tempo de espera entre dois consecutivos estágios é assumido como exponencial.

Posteriormente, em 1961, este modelo foi generalizado (9,48), derivando o “modelo de múltiplos estágios” de Armitage-Doll, que assume que um neoplasma maligno é originado de uma única célula-mãe normal suscetível que se transforma em totalmente maligna após sofrer uma série de mudanças irreversíveis (geralmente 6 ou 7 dependendo do tipo de câncer) devido a um componente hereditário e, para alguns tumores, também pela ação de um agente cancerígeno. Supõe-se que a taxa de transformação da i -ésima célula (r_i) varia linearmente com a exposição do agente cancerígeno na idade t do indivíduo [$r_i = \alpha_i + \beta_i f(t)$] (9,48). Este modelo tem se mostrado adequado para estimativas da incidência de câncer de estômago, pulmão e outros tumores em adultos que não dependem de fatores do tipo mudanças hormonais (como é o caso do câncer de mama) .

O modelo pode ser esquematizado da seguinte maneira:

onde:

C_0 : célula normal

C_k : células intermediárias , $k = 1, 2, \dots, (n-1)$

C_n : célula maligna

λ_k : taxas de transição do estágio C_k para o estágio C_{k+1} , $k = 0, 1, \dots, (n-1)$

$(\lambda_k)^{-1}$ (em anos): é a média de tempo no estágio C_k (as transformações celulares são instantâneas)

Dessa maneira o coeficiente de incidência $[I(t)]$ específico na idade t (por 100.000 habitantes) pode ser estimado por:

$$I(t) = N \lambda_0 \lambda_1 \dots \lambda_{n-1} t^{n-1} / (n-1)!$$

N : número total de células suscetíveis no tecido.

O modelo de múltiplos estágios de Armitage-Doll tem sido amplamente aceito, pois alguns pontos da história natural do câncer já estão definidos: é um processo de estágios múltiplos que ocorre através de um período de tempo variável, dependendo do tipo e localização do câncer e pode levar muitos anos até que o indivíduo morra por essa doença.

Outros autores sugeriram modificações a esse modelo, como FISHER (62) que propôs que o clone de uma célula pré-neoplásica epitelial se expande, lateralmente, como um disco circular a uma taxa constante e, dessa maneira, o raio dos clones é proporcional à idade de sua célula-mãe e sua área, proporcional ao quadrado de sua idade; ou como NEYMAN & SCOTT (62) que apresentaram um modelo de três estágios em que as células pré-neoplásicas cresciam de acordo com um modelo estocástico de nascimento e morte.

Outra aplicação deste modelo tem sido feita por diversos estudiosos que verificam o impacto da exposição a agentes cancerígenos na incidência de câncer, por meio de simulação.

A observação de dados de mortalidade e achados experimentais em animais já haviam demonstrado, em primeiro lugar, o grande período de latência entre a exposição ao cancerígeno e o desenvolvimento do tumor e, também, que a incidência de câncer tende a ser proporcional à concentração e ao tempo de exposição ao agente (3,63).

Diferentemente do que pode ser produzido em animais, o homem não está sob influência de cancerígenos em condições experimentais. A ação desses agentes pode ter ocorrido em diferentes períodos históricos, dependendo do indivíduo, e variar tanto de intensidade no tempo, quanto na idade do indivíduo afetado e a avaliação do efeito e sua interação com fatores genéticos e ambientais não pode ser feita diretamente de dados epidemiológicos, o que mostra a importância da simulação através de modelos matemáticos.

BROWN & CHU, 1989, (9) empregaram o modelo de Armitage-Doll para simular a ação de dois agentes atuando no mesmo estágio (efeito aditivo) ou em diferentes estágios (efeito multiplicativo) do processo cancerígeno e com tempos de exposição constantes ou não.

Já MAZUMDAR e colaboradores, 1989, (41) verificaram o efeito da exposição ao arsênico a que estavam sujeitos trabalhadores de uma fundição de cobre na incidência de câncer do pulmão. Simularam a ação do cancerígeno nas diferentes fases da carcinogênese em relação à idade do indivíduo quando do início da ação do agente, duração e tempo de exposição e tempo entre o término da exposição e o diagnóstico do câncer.

Em que pese a ampla aceitação do modelo de Armitage-Doll (1,9,10,22,37,39,48,49), MOOLGAVKAR & VENZON propuseram em 1976 (46) e, mais recentemente, entre 1986 e 1988 (22,49,62), aperfeiçoaram um modelo alternativo com a justificativa de que, experimentalmente, não foram comprovados mais que dois estágios na teoria da carcinogênese. Além disso, o modelo

de Armitage-Doll não contempla o crescimento, diferenciação e morte das células pré-neoplásicas; e, por último, o modelo não explica, satisfatoriamente, cânceres de origem unicamente hereditária como o retinoblastoma e o tumor de Wilms.

Apresentam, então, um modelo que, combinando componentes probabilísticos e determinísticos, explica a teoria da carcinogênese a partir da transformação e crescimento dos clones em dois estágios. Supõe que cada indivíduo tem um determinado número de células normais suscetíveis (consideradas variáveis aleatórias que obedecem a uma distribuição de Poisson) e que, em determinado momento, provavelmente devido à atuação de um gene (48), se transformam em células intermediárias que, após um período de tempo, se modificam ou em células intermediárias ou em malignas ou morrem. Esse processo é irreversível e ocorre de forma independente de um clone para outro. O esquema abaixo representa o processo descrito:

FIGURA 1

onde:

N : número de células normais suscetíveis

I : número de células intermediárias

M : número de células malignas

D : número de células mortas

$\beta_1(t)$: taxa de óbito de uma célula normal (por ano)

$\beta_2(t)$: taxa de óbito de uma célula intermediária (por ano)

$\alpha_1(t)$: taxa de divisão de uma célula normal em 2 células normais (por célula, por ano)

$\alpha_2(t)$: taxa de divisão de uma célula intermediária em 2 intermediárias (por célula, por ano)

$\mu_1(t)$: taxa da primeira mutação (por célula, por ano)

$\mu_2(t)$: taxa da segunda mutação (por célula, por ano)

$\theta(\Delta t)$: probabilidade de 1 ou mais eventos ocorrerem no intervalo de tempo Δt

$P_1 = \mu_1(t) \Delta t + \theta(\Delta t)$: probabilidade de uma célula normal se transformar em 2 células intermediárias no intervalo de tempo Δt .

$P_2 = \alpha_1(t) \Delta t + \theta(\Delta t)$: probabilidade de uma célula normal se transformar em 2 células normais no intervalo de tempo Δt .

$P_3 = \beta_1(t) \Delta t + \theta(\Delta t)$: probabilidade de uma célula normal morrer ou não se dividir no intervalo de tempo Δt .

$P_4 = \alpha_2(t) \Delta t + \theta(\Delta t)$: probabilidade de uma célula intermediária se transformar em 2 células intermediárias no intervalo de tempo Δt .

$P_5 = \beta_2(t) \Delta t + \theta(\Delta t)$: probabilidade de uma célula intermediária morrer ou não se dividir no intervalo de tempo Δt .

$P_6 = \mu_2(t) \Delta t + \theta(\Delta t)$: probabilidade de uma célula intermediária se transformar em 1 célula intermediária e outra maligna no intervalo de tempo Δt .

O conjunto de células geradas passa a ser detectável após atingir um determinado volume. A partir desse esquema demonstra-se que o coeficiente de incidência por câncer $[I(t)]$ (por 100.000 habitantes) é dado por:
$$I(t) = \mu_1 \mu_2 \int_0^t N(i) \exp [(\alpha_2 - \beta_2) (t-i)] di \quad i: \text{idade}$$

Este modelo explica, satisfatoriamente, a predisposição genética de alguns cânceres, principalmente aqueles com alta incidência na infância, as influências das variações hormonais na incidência do câncer de mama em mulheres e o aumento da incidência de alguns tipos de câncer a partir dos 40 anos. Contudo, as conclusões foram apenas teóricas e, em algumas situações, a estimação dos parâmetros foi feita por meio de experimentos em animais (principalmente ratos). Foi testado apenas para o câncer de mama em mulheres, quando foi considerado bom para estimar a sua incidência (47).

Também para este modelo foram feitas simulações sobre qual o efeito da exposição de agentes cancerígenos na divisão de células. Os agentes foram divididos em dois grupos, o primeiro formado pelos iniciadores quando atuam no clone e o outro formado por agentes promotores (22,49,62) ou inibidores (62) de tumor, estes dois últimos grupos de agentes atuando

nas células intermediárias. Estudou-se a sua influência na incidência de câncer quando a atuação ocorresse na primeira e/ou na segunda etapa do processo.

Modelos matemáticos que explicam as neoplasias por meio da teoria da carcinogênese têm sido utilizados com sucesso, como já descrito, porém muitas das suposições que foram feitas só puderam ser comprovadas através de experimentos em animais. Talvez, com o grande desenvolvimento na área da biologia molecular que tem ocorrido nos últimos tempos, essas suposições possam ser avaliadas com maior propriedade e novas descobertas possam ser facilmente incorporadas aos modelos. De acordo com MOOLGAVKAR (48), “... nós estamos ainda elaborando a melodia. Harmonia e contraponto devem vir mais tarde”.

Uma outra forma de estimar a incidência por câncer através de modelos matemáticos é analisar o processo mórbido sob outro prisma, utilizando os princípios do modelo “saúde-doença” elaborado por Chiang (12,50).

O primeiro modelo matemático a fazer uma abordagem mais genérica do processo saúde-doença foi o desenvolvido por CHIANG, 1965, (11,12,23,50), que idealizou um modelo estocástico, baseado no processo de Markov a tempo contínuo. Este é um modelo teórico, cujo principal objetivo era definir um “índice de saúde” único, variando de 0 a 100% e não foi testado, pois não dizia respeito a nenhuma doença especificadamente. Considerou-se que a saúde de um indivíduo é um fenômeno dinâmico, variando continuamente através do tempo. Desta maneira ela é formada por múltiplos estados de saúde ordenados que vão do completo bem estar até o extremo da doença, terminando com o óbito. Todos os estados são transitórios e o indivíduo pode, em determinado momento, estar em apenas um estado. Exceção a isso é o óbito, que é um estado absorvente (isto é, uma vez atingido esse compartimento a pessoa ali permanece). Este modelo ficou conhecido como “modelo de saúde-doença” e serviu de base para que outros estudiosos utilizassem esta teoria para doenças específicas como o câncer (39,64,65) e o infarto do miocárdio (13).

A partir do preceito de que a distribuição da morbidade na população é o primeiro determinante nos riscos de mortalidade, da mesma maneira que os efeitos da mortalidade definem os sobreviventes da doença (12,19,21,38,39,50), MANTON & STALLARD (39), KLEMENTIEV (64) e VERDECCHIA e colaboradores (65) propuseram novos modelos, utilizando a teoria do modelo de “saúde-doença” de Chiang.

MANTON & STALLARD, 1982, (39) propuseram um modelo bio-atuarial para estimar a incidência e a mortalidade anual por câncer de estômago em homens brancos americanos, utilizando um modelo estocástico de compartimentos. Este modelo foi construído a partir de uma série histórica de mortalidade de homens de 30 a 70 anos por um período de 28 anos (1950 a 1977), aplicando o modelo de crescimento de tumor de Armitage-Doll (utilizado como estado de transição de sadio para doente) e as probabilidades de morte de uma tábua de vida de múltiplo decremento

(utilizado como estado de transição de doente para óbito pela doença). O modelo foi considerado bom e recomendado para aquelas patologias em que se conhece o estado de transição. Infelizmente para se construir os parâmetros deste modelo é necessária uma série histórica de 28 anos de mortalidade, informação esta não disponível em nossa realidade.

VERDECCHIA e colaboradores, 1985, (64) testaram um modelo determinístico desenvolvido por KLEMENTIEV no *International Institute for Applied System Analysis (IIASA, Laxemburg, Austria)* para doenças crônico-degenerativas para populações estacionárias, utilizando dados do Registro de Câncer da Província de Varese (Itália). O modelo foi idealizado para populações estacionárias, com mortalidade, incidência e probabilidades de sobrevivência também estáveis, para doenças de caráter irreversível e assumindo que os coeficientes de mortalidade por outras causas, que não a doença considerada, são os mesmos para as pessoas doentes e as não doentes. O modelo foi considerado bom para os cânceres de algumas localizações, como estômago, pulmão e brônquios, colo e reto, porém ruim para bexiga urinária (homens) e mama (mulheres).

Mais recentemente, em 1989, o mesmo grupo, VERDECCHIA e colaboradores (65), publicou um trabalho mostrando um modelo estocástico desenvolvido por eles para estimar a incidência e a prevalência de doenças de caráter irreversível, na ausência de riscos competitivos, utilizando dados de população, mortalidade e tempo de sobrevivência após o diagnóstico. O modelo foi testado para câncer de pulmão e mama (mulheres) utilizando os dados de população e mortalidade da província de Varese (Itália) e as informações sobre sobrevivência do Registro de Câncer de Genebra (Suíça) e considerado bom para as estimativas de incidência. As estimativas sobre a prevalência não puderam ser avaliadas devido à inexistência de dados.

Embora estes modelos não utilizem a teoria de riscos competitivos, os resultados foram muito satisfatórios quando comparados com dados reais e o seu grande valor consiste no emprego de séries históricas de mortalidade, visto que as séries existem, são de fácil acesso e, em especial no caso do câncer, as regras de seleção de causa básica sempre indicam o neoplasma maligno como elemento iniciador do processo mórbido (16,37,54,65). Isto propicia que o processo de saúde-doença-óbito seja estudado de forma mais abrangente, dentro do tempo.

Nos últimos anos têm surgido novos modelos para determinar a incidência de câncer. ROBINS & GREENLAND, 1989, (55) propuseram um modelo estocástico para determinar a probabilidade de causação de um fator de risco para doenças crônicas em uma coorte, supondo que a probabilidade pode variar de um indivíduo para outro devido à interação do agente com fatores genéticos. Nesse modelo foi incorporada a teoria de dados censurados e riscos competitivos e foi aconselhada a sua utilização para o câncer.

Já KAMPERT & SILVERS, 1989, (34) desenvolveram um teste para verificar a adequação de modelos multifásicos de câncer para parâmetros de transição que obedecem à

distribuição de Poisson e são dependentes da idade e da exposição a agentes cancerígenos.

Em outro trabalho, ANDERSON & ROLLER, 1991, (1) discutem pontos de equilíbrio em modelos de compartimentos não lineares e fazem simulações de algumas situações, derivando estados de estabilidade e tempos de residência nos compartimentos.

CALDERÓN & KWEMBE, 1991, (10) discutem um modelo de crescimento de tumor baseado na lei de Gompertz e verificam o efeito da entropia no mecanismo.

Pode-se observar que há diversos tratamentos dados aos modelos matemáticos que estimam a incidência de câncer e até o presente, ainda se discute um modelo adequado, pois a carcinogênese abarca um intrincado processo que envolve fatores genéticos, hereditários e ambientais difíceis de serem traduzidos de maneira completa e simples por fórmulas matemáticas. Isso mostra que há muito o que se pesquisar e elaborar nessa área .

Todos os estudos apresentados abriram novas perspectivas para os estudos de morbidade em câncer. Em nosso meio, a carência de dados e de recursos nesta área, aliada às dificuldades das fontes de dados de morbidade existentes, fazem com que a busca de métodos indiretos para se medir morbidade seja extremamente oportuna.

No Brasil há diversos pesquisadores que desenvolvem modelos matemáticos para estimar a incidência de algumas doenças, porém a ênfase tem sido dada às doenças transmissíveis. Nada há publicado por pesquisadores brasileiros, até o presente momento, em relação a métodos indiretos para estimar a incidência de doenças crônico-degenerativas.

No caso particular das neoplasias malignas há diversos modelos desenvolvidos e testados no exterior, porém não foi verificada a adequação dos mesmos aos dados de incidência existentes em nosso meio. Sendo assim, é importante, pois, que se desenvolva uma linha de pesquisa em modelos matemáticos para doenças não transmissíveis no Brasil, onde a carência de dados de morbidade só aumenta as dificuldades de um planejamento eficaz em saúde.

II - MODELOS

1. MODELO KLEMENTIEV (MODELO 1) (64)

Este é um modelo determinístico para estimar os coeficientes de incidência para doenças crônico-degenerativas para populações estacionárias e testado, para cada um dos sexos separadamente, utilizando dados de câncer do estômago .

1.a. Suposições

- . a população é estacionária e estável;
- . a mortalidade, a incidência e as taxas de sobrevivência para a doença são estáveis;
- . a doença tem caráter irreversível;
- . os indivíduos doentes não morrem pela causa considerada durante o primeiro ano de doentes (isto é, o coeficiente de mortalidade pela doença no primeiro ano é zero);

1.b. O Modelo

Sejam:

P_i : número de indivíduos na i-ésima faixa etária

h_i : número de indivíduos sadios na i-ésima faixa etária

m_i : coeficiente específico de incidência na i-ésima faixa etária

D_i^T : coeficiente de mortalidade total específico para a i-ésima faixa etária

D_i^* : coeficiente de mortalidade específico para a i-ésima faixa etária e para a causa considerada

D_i : coeficiente de mortalidade por outras causas que não a considerada específico para a i-ésima faixa etária

d_{ij} : coeficiente de mortalidade para indivíduos doentes específico para a i-ésima faixa etária, para indivíduos que contraíram a doença na faixa etária j

S_{ij} : taxa de sobrevivência de j anos para indivíduos que contraíram a doença na i-ésima faixa etária

b_{ij} : coeficiente de mortalidade total para os indivíduos que ficaram doentes na faixa etária j específico para a i-ésima faixa etária

Tem-se que:

$$d_{ij} = (S_{i,j-1} - S_{ij}) / S_{i,j-1}$$

$$D_i = D_i^T D_i^*$$

$$b_{ij} = D_{i+j} + d_{ij}$$

$$a_{ij} = \prod_{k=1}^j (1 - b_{ik}^i)$$

Com estas equações pode-se definir a população total como sendo

$$P_i = \underbrace{h_i}_{\text{indiv. sadios}} + \underbrace{m_{i-1} h_{i-1}}_{\text{indiv. que contraíram doença há 1 ano atrás}} + \underbrace{\sum_{j=2}^{i-1} m_{i-j} h_{i-j} a_{i-j,j-1}}_{F_i = \text{indiv. que contraíram a doença à 2 ou + anos atrás}} \quad (1)$$

Ao se multiplicar os dois lados da igualdade pelos respectivos coeficientes de mortalidade total, obtém-se

$$P_i D_i^t = \underbrace{h_i D_i + m_{i-1} h_{i-1} b_{i-1,1} + \sum_{j=2}^{i-1} m_{i-j} h_{i-j} a_{i-j,j-1} b_{i-j,1}}_{G_i} \quad (2)$$

Sob as suposições anteriormente citadas o sistema formado pelas equações (1) e (2) pode ser resolvido em função de m_i e h_i , chegando-se às seguintes equações:

$$m_{i-1} = \frac{P_i (D_i^t - D_i) - G_i + F_i D_i}{h_{i-1} (b_{i-1,1} - D_i)} \quad (3)$$

$$h_i = \frac{P_i (D_i^t - b_{i-1,1}) - G_i + F_i b_{i-j,1}}{D_i - b_{i-1,1}} \quad (4)$$

1.c. Comentários

Na publicação que descreve este modelo há algumas suposições que não estão claras, talvez pelo fato de que este modelo foi desenvolvido na Áustria e o artigo é de pesquisadores italianos que testaram a validade do mesmo com dados de câncer da província de Varese (Itália).

Foi necessário, então, que fosse definida aqui uma hipótese inicial: não existem indivíduos doentes com menos de 25 anos. Isso significa que $h_0 = P_0$, $m_0 = 0$ e $D_0^* = 0$.

Foi escolhido o câncer de estômago para se testar este modelo porque foi uma das localizações em que o modelo foi mais preciso quando testado com dados italianos e é uma das localizações com maior incidência no Município de São Paulo, para ambos os sexos.

2. MODELO VERDECCHIA (MODELO 2) (65)

Este é um modelo probabilístico para estimar a incidência e a prevalência de doenças crônico-degenerativas em populações não estáveis. Foi testado para o câncer de pulmão (ambos os sexos) e o de mama feminina.

2.a. Suposições

- . a população não precisa ser estável;
- . a doença é irreversível;
- . o grupo de pessoas doentes tem a mesma distribuição etária que os não doentes;
- . não existem riscos competitivos;
- . tanto o diagnóstico quanto o óbito pela doença ocorrem sempre no ponto médio entre dois aniversários consecutivos, exceto para aqueles indivíduos que ficam doentes e morrem dentro do mesmo ano. Para estes últimos é assumido o tempo médio da doença de seis meses.

2.b. O Modelo

O esquema abaixo descreve o processo mórbido:

FIGURA 2

onde

$\mu(x)$ = probabilidade de ser diagnosticado um caso novo entre as idades x e $(x+1)$

$\alpha(x)$ = probabilidade de morrer pelas outras causas à idade x

$\gamma(x)$ = probabilidade de morrer pela causa específica à idade x

$\beta(x,y)$ = risco de morrer por outra causa à idade x para pessoas que ficaram doentes à idade y
 $\delta(x,y)$ = risco de morrer pela causa específica à idade x para pessoas que ficaram doentes à idade y

Para a elaboração do modelo os autores consideraram que $\alpha(x)$, $\beta(x,y)$, $\delta(x,y)$ e $\gamma(x)$ são parâmetros de transição conhecidos e que o único parâmetro a ser estimado é $\mu(x)$.

Definem-se, ainda, mais duas variáveis que serão utilizadas no modelo:

$S(x)$: probabilidade de sobrevivência para uma coorte de idade, à idade x

$v(x)$: probabilidade de estar doente à idade x , condicional à sobrevivência até esta idade

$$S(x) = \exp \left[- \int_0^x \alpha(u) du \right] \quad (1)$$

$$\gamma(x) = \int_a^x [1 - v(a)] \mu(a) \delta(x, a) \exp \left\{ - \int_a^x [\beta(u,a) - \alpha(u)] du \right\} da \quad (2)$$

$$v(x) = \int_a^x [1 - v(a)] \mu(a) \exp \left\{ - \int_a^x [\beta(u,a) - \alpha(u)] du \right\} da \quad (3)$$

Os parâmetros foram definidos, primeiramente, em termos contínuos e as equações derivadas da mesma maneira. Porém os autores justificaram que, como os dados disponíveis são sempre discretos, transformaram as equações em termos discretos, conforme apresentadas a seguir:

$$\gamma_i = \underbrace{(1-v_i) \mu_i d_{ii}}_{\text{pessoas que contraem a doença e morrem dentro do mesmo ano}} + \sum_{j=0}^{i-1} (1-v_j) \mu_j d_{ij} S_{ij} \quad (4)$$

$$v_i = \sum_{j=0}^i (1-v_j) \mu_j S_{ij} \quad (5)$$

$$V_0 = 0$$

E os parâmetros passam a ser definidos como:

v_i = proporção de indivíduos doentes à i -ésima faixa etária

μ_i = probabilidade de ser diagnosticado um caso novo à i -ésima faixa etária em indivíduos que estavam sadios até então

d_{ij} = probabilidade de morrer por uma específica doença à i -ésima faixa etária para indivíduos

doentes que contraíram a doença na j -ésima faixa etária

γ_i = probabilidade de morrer por uma específica doença à i -ésima faixa etária para todas as pessoas que estavam vivas à essa faixa etária

S_{ij} = probabilidade de sobrevivência à i -ésima faixa etária para indivíduos doentes que contraíram a doença há j anos atrás.

Como não se dispunha da informação acerca de d_{ij} , ele foi utilizado como uma função de S_{ij} . Assim, $d_{ij} = 1 - S_{ij}$.

Substituindo os valores de γ_i , d_{ij} e S_{ij} , que são conhecidos, nas equações (4) e (5) obtém-se os valores de ν_i e μ_i . Multiplicando-se os valores dos μ_i 's pela população da faixa etária correspondente obtém-se o número de casos novos daquela faixa etária.

2.c. Comentários

Este modelo foi proposto pelos autores como um modelo que se adequava tanto aos cânceres com um comportamento regular com a idade e, por isso, fáceis de serem modelados, como o de pulmão, quanto a outros de comportamento irregular, como o de mama, que é função não tanto da idade, mas, principalmente, de fatores reprodutivos e alterações hormonais.

Os autores trabalharam com probabilidade de sobrevivência constante para todas as idades devido à não disponibilidade de dados de sobrevivência segundo faixa etária. Neste trabalho, as taxas de sobrevivência utilizadas não foram constantes como pode ser observado nos Anexos 11 e 12.

Também consideraram que seria melhor utilizar-se a regressão logística em função da idade e coorte de nascimento como estimador de máxima verossimilhança para a incidência para não trabalhar com dados de mortalidade e evitar, assim, os problemas inerentes a eles (como o número pequeno de casos, grande oscilação nos dados e outros). Porém para este trabalho optou-se por utilizar as equações (4) e (5) para se calcular os parâmetros do modelo por considerar, em primeiro lugar, que estas são as equações que descrevem o processo mórbido de acordo com a idealização dos autores e, por outro lado, não ficou clara, na publicação, a função de máxima verossimilhança escolhida (se ela é função da incidência ou da mortalidade).

3. MODELO MANTON (MODELO 3) (39)

Este é um modelo probabilístico para estimar a incidência e a mortalidade de cânceres que ocorrem principalmente na idade adulta e testado pelos autores para câncer de estômago em homens brancos americanos.

3.a. Suposições

- . a doença tem caráter irreversível;
- . o câncer ocorre principalmente na idade adulta, com a incidência nas idades abaixo de 30 anos tendendo a zero;
- . não existem riscos competitivos;
- . as causas de óbito são independentes;
- . os indivíduos doentes não morrem pela doença durante o primeiro ano de diagnóstico;
- . toda a população faz aniversário em primeiro de janeiro.

3.b. O Modelo

O esquema abaixo representa o processo mórbido deste modelo:

FIGURA 3

$\lambda_1(a)$ representa o parâmetro de transição de sadio para doente e a função de máxima verossimilhança escolhida para estimá-lo baseou-se no modelo de “múltiplos estágios de Armitage-Doll”:

$$\lambda_1(a_i) = \frac{\alpha (a_i - t)^{m-1} m s}{t + \alpha (a_i - t)^m}$$

onde:

m = número de mutações necessárias para o tumor adquirir volume detectável

$\bar{\alpha}$ = produto das probabilidades das m mutações

a_i = idade do início do tumor

t = tempo médio de progressão da doença, isto é, o tempo entre a idade em que o tumor iniciou o seu crescimento e o óbito do indivíduo

s = volume mínimo detectável do tumor

O parâmetro de transição de doente para óbito por câncer, $\lambda_2(a)$ foi estimado utilizando a técnica da tabela de vida de múltiplo decremento com algumas modificações, pois leva em conta, também, a duração do crescimento do tumor.

3.c. Comentários

Este modelo não foi testado por este trabalho, pois necessita-se de uma série histórica de mortalidade de 28 anos e somente a partir de 1970 as estatísticas de mortalidade, para o Estado de São Paulo, começaram a ser publicadas por local de residência. Por isso não dispõe-se dessa série de mortalidade para o Município de São Paulo, o que inviabilizou a análise do modelo.

Porém de todos os modelos citados foi o que melhor predisse os valores da incidência de câncer (no caso, o câncer de estômago) e o único que incorpora conhecimentos a respeito da teoria da carcinogênese, aliado a dados de sobrevivência, sendo estes os motivos pelos quais considerou-se importante a descrição deste modelo com mais detalhes.

4. MODELO PROPOSTO (MODELO 4)

Este é um modelo probabilístico genérico que pode ser aplicado a todas as doenças que possuam as seguintes características:

- . terem caráter irreversível, isto é, não existir cura;
- . a incidência, a mortalidade, e as taxas de sobrevivência são constantes, pelo menos, durante um ano.

Este modelo foi testado para os cânceres das localizações utilizadas pelos outros modelos (1 e 2), ou seja, estômago, traquéia, brônquios e pulmão e mama feminina.

A observação dos dados de mortalidade por câncer mostra que, para algumas localizações, a mortalidade aumenta rapidamente com a idade e, para outras o mesmo fenômeno não ocorre (3,18). Além disso, existe um consenso de que pessoas de diferentes idades e sexos têm desiguais suscetibilidades às doenças e as doenças agem diferentemente nelas. Em quase todas as localizações, os coeficientes de incidência de câncer são maiores para o sexo masculino do que para o feminino. Portanto para se descrever adequadamente o processo mórbido, o modelo deve considerar a população dividida em sub-populações de acordo com estas variáveis (12,18,21,23,50). Dessa maneira, o modelo foi elaborado considerando a população dividida em faixas etárias de 5 anos, dos 25 aos 69 anos e, para cada localização, o modelo foi testado para cada sexo separadamente, com exceção do câncer de mama que foi feito apenas para o sexo feminino.

4.a. Suposições

- . a doença é irreversível;
- . não existem riscos competitivos;
- . a população é estável;
- . a incidência, a mortalidade e as taxas de sobrevivência são constantes durante o ano.

4.b. O Modelo

A Figura 4, a seguir, representa a descrição do processo mórbido que foi assumido pelo modelo:

FIGURA 4

onde

$\mu_1(x_1)$ = parâmetro de transição de sadio para doente à idade x_1 , isto é, o risco instantâneo de um indivíduo passar do estado de sadio para o de doente à idade x_1

$\mu_2(x)$ = parâmetro de transição de sadio para óbito por outras causas à idade x , isto é, o risco instantâneo de um indivíduo passar do estado de sadio para o de óbito por outras causas à idade x

$\mu_2(x_1+a)$ = parâmetro de transição de doente para óbito por outras causas para indivíduos que adquiriram câncer à idade x_1 e morreram “a” anos depois, isto é, o risco instantâneo de um indivíduo passar do estado de doente para o de óbito por outras causas à idade (x_1+a)

$\mu_2(x) = \mu_2(x_1+a) \iff x = x_1+a$ (não existem riscos competitivos)

$\mu_3(x_1+a)$ = parâmetro de transição de doente para óbito por câncer da localização considerada para indivíduos que adquiriram o câncer à idade x_1 e morreram “a” anos depois, isto é, o risco instantâneo de um indivíduo passar do estado de doente para o de óbito por câncer da localização considerada à idade (x_1+a)

$\mu_4(x)$ = probabilidade de permanecer no estado transitório de sadio

$\mu_5(x)$ = probabilidade de permanecer no estado transitório de doente

Sabendo-se que os riscos instantâneos $\mu_i(x)$ se mantêm constantes durante o ano, os de mortalidade podem ser aproximados pela probabilidade de óbito da causa considerada.

Seja $p_{i,j}$ a probabilidade de sobrevivência pela causa i à idade j . Essa probabilidade é a taxa de sobrevivência relativa para um ano. Dessa maneira a probabilidade de morte pela causa

i à idade j ($q_{i,j}$) pode ser calculada como:

$$q_{i,j} = 1 - p_{i,j}$$

Desse modo podemos definir os riscos instantâneos de morte $[\mu_2(x) \text{ e } \mu_3(x)]$ em função das probabilidades de morte ($q_{i,j}$):

$$\mu_2(x) = q_{2,x}$$

$$\mu_3(x) = q_{3,x}$$

Para se elaborar o modelo, considera-se uma coorte de pessoas que no início do ano estão todas no estado inicial “sadio”. Esta coorte será denominada de P_0 .

Esta população inicial (P_0) está sujeita, durante o ano, a dois riscos: $\mu_1(x)$ e $q_2(x)$, o que significa que ao final de um ano parte destes indivíduos permanecerão no estado de “sadios” (P_S), outra parte passará para o estado absorvente de “óbito por outra causa que não o câncer considerado” e uma última parte terá passado para o estado transitório de “doente” (P_D), isto é, serão pacientes portadores do câncer da localização considerada.

Assume-se que a população de doentes do meio do período será uma estimativa do número de casos novos durante o ano, pois a população de doentes é crescente durante o ano.

E este é o objetivo do modelo: determinar qual a proporção de P_0 que, no meio do ano, estará doente.

Os indivíduos doentes (P_D), como não podem se curar, estão sujeitos a dois riscos de morte: um pela doença em questão ($q_{3,x}$) e outro por outras causas ($q_{2,x}$).

Logo, após um ano, parte dos indivíduos doentes permanecerá doente, outra parte terá morrido por outra causa que não o câncer da localização considerada ($O_{O,D}$) e o restante terá morrido por câncer dessa localização (O_C).

Vale lembrar que o risco de morrer por outras causas para indivíduos doentes é, supostamente, o mesmo que para não doentes.

Para estabelecer-se as equações do sistema apresentado na Figura 4 serão definidos os seguintes parâmetros:

- $P_{0,x}$: população inicial da coorte de indivíduos sadios na idade x = População de vivos
- $\bar{P}_{0,x}$: população de sadios no meio do ano na idade x
- $P_{D,x}$: população de doentes ao fim de um ano na idade x

- $\bar{P}_{D,x}$: população de doentes no meio do ano na idade x
 $P_{S,x}$: população de sadios ao fim de um ano na idade x
 $O_{i,x}$: número de óbitos por todas as causas na idade x
 $O_{C,x}$: número de óbitos por câncer da localização considerada na idade x
 $O_{O,x}$: número de óbitos por outras causas que não o câncer da localização considerada na idade x
 $O_{O,S,x}$: número de óbitos por outras causas que não o câncer da localização considerada de indivíduos sadios que morreram na idade x
 $O_{O,D,x}$: número de óbitos por outras causas que não o câncer da localização considerada de indivíduos doentes que morreram na idade x
 $q_{i,x}$: probabilidade de morte pela causa i à idade x
 $q_{o,x}$: probabilidade de morte por outras causas que não i à idade x
 $q_{o,S,x}$: probabilidade de morte por outras causas que não i à idade x para indivíduos sadios
 $q_{o,D,x}$: probabilidade de morte por outras causas que não i à idade x para indivíduos doentes

A população de doentes no meio do ano é representada por:

$$\bar{P}_{D,x} = P_{D,x} + 0,5 O_{C,x} + 0,5 O_{O,D,x} \quad (1)$$

O número de óbitos por outras causas que não o câncer da localização considerada pode ter ocorrido tanto por indivíduos doentes como por não doentes, ou seja:

$$O_{O,x} = O_{O,S,x} + O_{O,D,x} \quad (2)$$

E, qualquer que seja a causa do óbito, este número é determinado pela população da faixa etária considerada multiplicada pelo risco de morte da causa, isto é:

$$O_{C,x} = P_{D,x} q_{i,x} \Rightarrow P_{D,x} = O_{C,x} / q_{i,x} \quad (3)$$

$$O_{O,S,x} = P_{S,x} q_{o,s,x} \quad (4)$$

$$O_{O,D,x} = P_{D,x} q_{o,D,x} \quad (5)$$

Substituindo-se as equações (4) e (5) na equação (2) tem-se que:

$$O_{O,x} = P_{S,x} q_{o,s,x} + P_{D,x} q_{o,D,x} \quad (6)$$

Como não existem riscos competitivos (isto é, $q_{o,s,x} = q_{o,D,x} = q_{o,x}$) pode-se reescrever a equação (6) da seguinte maneira:

$$O_{O,x} = (P_{S,x} + P_{D,x}) q_{o,x} = P_{0,x} q_{o,x} \Rightarrow q_{o,x} = O_{O,x} / P_{0,x} \quad (7)$$

Substituindo-se a equação (7) na equação (5) pode-se reescrever o número de óbitos por outras causas que não o câncer da localização considerada para indivíduos doentes como sendo:

$$O_{O,D,x} = P_{D,x} \frac{O_{O,x}}{P_{0,x}} \quad (8)$$

Substituindo-se a equação (3) na equação (8) tem-se que:

$$O_{O,D,x} = \frac{O_{C,x}}{q_{i,x}} \frac{O_{O,x}}{P_{0,x}} \quad (9)$$

Finalmente, ao se substituir as equações (3) e (9) na equação (1) consegue-se que a população de doentes no meio do ano possa ser reescrita apenas em função da população, do número de óbitos e da probabilidade de morte pelo câncer da localização considerada:

$$\begin{aligned} \bar{P}_{D,x} &= \frac{O_{C,x}}{q_{i,x}} + 0,5 O_{C,x} + 0,5 \frac{O_{C,x}}{q_{i,x}} \frac{O_{O,x}}{P_{0,x}} \Rightarrow \\ \bar{P}_{D,x} &= \left(\frac{1}{q_{i,x}} + 0,5 + \frac{O_{O,x}}{2 q_{i,x} P_{0,x}} \right) O_{C,x} \quad (10) \end{aligned}$$

A equação (10) fornece a estimativa do número de casos novos de câncer da localização considerada para um ano.

4.c. Comentários

Deve-se notar que para o presente trabalho o modelo é utilizado de maneira semelhante às aplicações usuais das tábuas de sobrevivência: o modelo é probabilístico na sua formulação e determinístico na sua aplicação. Não se incorporou, por exemplo, o cálculo das variabilidades dos estimadores nem de suas principais características. Embora se reconheça a relevância desses tópicos, optou-se por cuidar dessa abordagem em trabalho posterior.

III. OBJETIVOS

1. Verificar a adequação de alguns modelos existentes na literatura para estimação da incidência por câncer a partir da mortalidade, utilizando dados do Município de São Paulo, para o ano de 1978.
2. Propor um modelo probabilístico para estimar a incidência por câncer, a partir de dados de mortalidade.
3. Comparar os resultados obtidos pelos modelos propostos em literatura e os obtidos pelo modelo desenvolvido neste trabalho, com os dados de câncer coletados no Registro de Câncer do Município de São Paulo, referente ao ano de 1978.

IV. MATERIAL E MÉTODOS

Foram testados dois modelos encontrados na literatura: o desenvolvido por KLEMENTIEV (64) (Modelo 1) que utilizou dados de câncer de estômago e o de VERDECCHIA e colaboradores (65) (Modelo 2) que trabalhou com os dados de câncer de traquéia, brônquios e pulmão e mama feminina.

Além desses, foi desenvolvido e testado outro modelo (Modelo 4) para as localizações: estômago, traquéia, brônquios e pulmão e mama feminina.

Para a codificação dos tumores foi utilizada a Classificação Internacional de Doenças - 9ª revisão, 1975, (16) sendo consideradas as seguintes categorias:

151: Neoplasma maligno de estômago

162: Neoplasma maligno de traquéia, brônquios e pulmão

174: Neoplasma maligno de mama feminina

Para o exame dos Modelos 1, 2 e 4 foram empregados dados de população, óbito e taxas de sobrevivência por câncer, segundo sexo e idade para as localizações citadas anteriormente e, após a obtenção das curvas de incidência estimadas, foram utilizados os dados de incidência de câncer do Registro de Câncer de São Paulo referentes ao ano de 1978 para se comparar os resultados. Foi utilizado o ano de 1978, pois este é o último ano com dados de incidência publicados para o Município de São Paulo.

Para todos os modelos optou-se por trabalhar com a idade a partir de 25 anos, pois anteriormente à essa idade há poucos casos de câncer para as localizações escolhidas. A última faixa etária foi a de 65 a 69 anos, pois é sabido que nas idades mais avançadas há muitos erros na Declaração de Óbito no que se refere à informação da causa básica do óbito (entre outros itens). Diversos autores recomendam que se utilize a idade dos 20 ou 25 anos até os 70 ou 75 anos quando se quer testar modelos em câncer.

1. População

Os dados de população para o ano de 1978 foram retirados da publicação “Cancer Mortality in São Paulo-Brazil” (45). Esta população foi estimada através de interpolação linear das populações dos censos de 1970 e 1980 e está apresentada no Anexo 1.

2. Óbitos e Coeficientes de Mortalidade (36)

Os óbitos de residentes no Município de São Paulo, segundo sexo e idade foram fornecidos pela Fundação SEADE (Anexo 2).

Os óbitos por câncer e os coeficientes de mortalidade por câncer foram retirados do livro “Cancer Mortality in São Paulo-Brazil” (45).

Para o Modelo 1 utilizou-se os coeficientes específicos de mortalidade segundo sexo e idade (Anexo 3) para residentes no Município de São Paulo, 1978 e os coeficientes específicos de mortalidade por câncer de estômago, também segundo sexo e idade (Anexo 4).

Ao se analisar o Modelo 2 foram utilizados os coeficientes específicos de mortalidade por câncer de traquéia, brônquios e pulmão, segundo sexo e idade, (Anexo 5) e mama feminina (Anexo 6).

Já para o Modelo 4 necessitou-se não só dos coeficientes de mortalidade, mas, também, do número de óbitos gerais (Anexo 2), segundo sexo e idade, além dos de câncer de estômago (Anexo 7), traquéia, brônquios e pulmão (Anexo 8) e mama feminina (Anexo 9), também por sexo e idade.

3. Taxas de sobrevivência (63)

As taxas de sobrevivência utilizadas neste trabalho referem-se aos casos registrados no Registro de Câncer da Noruega durante o período de 1953 a 1967.

Após pesquisa em literatura, conseguiu-se informações sobre sobrevivência de câncer, segundo sexo e idade para as seguintes localidades: Noruega (1953-1967), Finlândia (1953-1974), Estados Unidos (registros hospitalares para o período de 1955 a 1964) e Inglaterra (Registro de Câncer de Birmingham, 1960 à, presumivelmente, 1970). Optou-se pelos dados da Noruega pelo fato de serem de base populacional e calculados por faixa etária. A defasagem entre o ano utilizado neste trabalho (1978) e o período do estudo da Noruega (1953-1967) foi considerado conveniente, pois se fossem utilizados dados referentes ao mesmo período, sendo a Noruega um país mais desenvolvido do que o Município de São Paulo, provavelmente as taxas de sobrevivência estariam super-estimadas.

O Modelo 1 utilizou taxas de sobrevivência relativas de 5 anos para câncer de estômago (Anexo 10) e o Modelo 2 taxas de sobrevivência relativas de 5 anos para câncer de traquéia, brônquios e pulmão (Anexo 11) e mama feminina (Anexo 12). O Modelo 4 utilizou as taxas de sobrevivência relativas de um ano para as mesmas localizações que os outros dois modelos (Anexos 10 a 12).

4. Coeficientes de Incidência de Câncer (36)

Os coeficientes de incidência para as localizações estudadas (estômago, traquéia, brônquios e pulmão e mama feminina) foram obtidos a partir dos casos registrados no Registro de Câncer de São Paulo em 1978 (anexos 13 a 15) e publicados no livro “Cancer Incidence in São Paulo County - Brazil” (44).

Vale ressaltar o fato de que estes dados foram trabalhados pelo Registro de Câncer de São Paulo manualmente e, por isso, sujeitos a erros. Somente no presente momento as fichas de coleta de dados estão sendo codificadas e digitadas de maneira mais criteriosa. Não há publicação sobre a fidedignidade dos mesmos e nem se conhece a sua precisão.

5. Análise

Foram comparados os resultados obtidos pelos modelos propostos em literatura (Modelos 1 e 2) e os obtidos pelo modelo proposto por este trabalho (Modelo 4), com os dados de câncer, para o Município de São Paulo, referente ao ano de 1978.

A comparação entre os resultados obtidos pelos modelos e os dados reais de incidência foi feita graficamente, em primeira instância e, após, realizou-se a análise das diferenças entre as curvas de incidência obtidas, através do teste χ^2 de Aderência (42). O χ^2 foi utilizado não só como medida de aderência das curvas, mas, também, para verificar as diferenças encontradas em cada faixa etária.

V. RESULTADOS E DISCUSSÃO

Na Tabela 1, a seguir, estão apresentados os resultados dos coeficientes de incidência para o câncer de estômago (sexo masculino) obtidos pelos modelos 1 e 4, bem como os coeficientes de incidência reais. Como pode ser observado, o Modelo 1 fornece estimativas razoáveis da incidência de câncer de estômago para o sexo masculino até a faixa etária de 45 a 49 anos, mostrando bom ajuste das estimativas até essa idade, porém a partir dos 50 anos, as estimativas ficam sub-enumeradas, com as diferenças entre os coeficientes estimados e os observados aumentando com a idade. O χ^2 de aderência foi de 270,3, significativo ($p < 0,001$) mostrando que a curva do Modelo 1 não se ajusta aos dados observados. Até a faixa etária dos 40 a 44 anos as diferenças entre as incidências observada e esperada são pequenas mostrando bom ajuste das curvas até essa idade.

Tabela 1 - Coeficientes de Incidência observado e estimados de câncer de estômago no sexo masculino (por 100.000 homens), segundo idade, modelos e valores dos desvios quadráticos (D.Q.*) do teste de aderência pelo χ^2 . Município de São Paulo, 1978.

IDADE	Coeficiente ^{1,2} Observado	Coef.estimado ¹ Modelo 1	D.Q.	Coef.estimado ¹ Modelo 4	D.Q.
25 - 29	1.2	0.0	1.2	0.0	1.2
30 - 34	2.6	5.5	3.2	3.1	0.1
35 - 39	12.2	11.1	0.1	18.0	2.8
40 - 44	34.5	21.7	4.7	35.3	0.0
45 - 49	50.1	35.0	4.6	72.2	9.7
50 - 54	107.5	45.6	35.6	109.0	0.0
55 - 59	139.5	55.1	51.0	153.6	1.4
60 - 64	191.6	90.6	53.2	198.0	0.2
65 - 69	310.0	119.8	116.7	308.0	0.0

1 : por 100 000 homens

2 : dados referentes ao Registro de Câncer de São Paulo (Anexo 13)

$$(O_i - E_i)^2$$

$$* : D.Q. = \frac{(O_i - E_i)^2}{E_i}$$

Já o Modelo 4 produz estimativas bem próximas aos valores reais, com um χ^2 de aderência de 15,4 ($p < 0,10$), o que mostra não haver diferenças significantes entre os valores estimados e os reais. Apenas para a faixa etária dos 45 a 49 anos houve um erro maior e o coeficiente estimado ficou cerca de 40% maior do que o observado.

O Gráfico 1 representa os dados referentes à Tabela 1 onde se percebe o melhor ajuste das estimativas dos coeficientes de incidência de câncer de estômago para o sexo masculino feitas pelo Modelo 4 em relação às feitas pelo Modelo 1.

Pode-se observar que até a idade de 44 anos as estimativas feitas por ambos os modelos são muito próximas dos valores reais e só a partir dessa idade é que as diferenças apresentadas pelo Modelo 1 passam a ser significantes.

GRÁFICO 1

A grande divergência na elaboração dos Modelos 1 e 4 é que o Modelo 1 acompanha uma coorte de indivíduos sadios até sua morte durante vários anos e, conseqüentemente, utiliza taxas de sobrevivência relativas acumuladas para cada coorte de idade. Sabe-se que as taxas de sobrevivência para o câncer podem não ser constantes durante muitos anos devido aos avanços nas técnicas de diagnóstico e terapia. Isto provavelmente estaria influenciando as estimativas feitas por este modelo para as idades mais avançadas, onde o efeito cumulativo das taxas de sobrevivência pode ser mais sentido.

Porém o Modelo 4 ao acompanhar uma coorte de indivíduos sadios por um ano não sofre o efeito cumulativo das taxas de sobrevivência relativa e a suposição de que ela seja constante por um ano é plausível. Talvez esta seja a explicação do porquê o Modelo 1 passa a não fornecer boas estimativas a partir dos 45 anos e o Modelo 4 continua sendo um bom modelo.

A comparação visual entre as curvas de incidência de câncer de acordo com a idade também pode ser feita através de gráficos onde os coeficientes de incidência aparecem em escala logarítmica. Este gráfico permite perceber as mudanças das curvas em termos relativos com maior clareza do que quando as incidências aparecem em escala aritmética, como no Gráfico 1, devido à grande variabilidade de seus valores.

Ao observar o Gráfico 2 onde as curvas de incidência aparecem em escala logarítmica, nota-se que as três curvas têm o mesmo formato. A curva do Modelo 4 está sempre oscilando entre a curva com os valores observados e as estimativas do Modelo 1 estão sempre sub-enumeradas.

É importante salientar que quando se deseja estudar tendências pode-se utilizar qualquer uma das curvas, pois em todas pode ser verificado que a incidência de câncer de estômago para o sexo masculino sempre é crescente com a idade.

GRÁFICO 2

A Tabela 2 apresenta os resultados das estimativas dos coeficientes de incidência de câncer de estômago para o sexo feminino. Pode-se observar, em relação ao Modelo 1, fenômeno semelhante ao que ocorreu com o sexo masculino: a partir dos 50 anos este modelo passa a não fornecer boas estimativas. Embora as diferenças sejam menos significantes, as estimativas para o sexo feminino também são sub-enumeradas mostrando um χ^2 de aderência de 102,5 ($p < 0,001$).

Ao se analisar os resultados do Modelo 4 percebe-se o bom ajuste aos dados reais, com um χ^2 de aderência de 8,6 ($p < 0,50$). Isso demonstra que não há diferenças estatisticamente significantes entre os valores observados e os estimados pelo Modelo 4.

Tabela 2 - Coeficientes de Incidência observado e estimados de câncer de estômago no sexo feminino (por 100.000 mulheres), segundo idade, modelos e valores dos desvios quadráticos (D.Q.*) do teste de aderência pelo χ^2 . Município de São Paulo, 1978.

IDADE	Coeficiente ^{1,2} Observado	Coef.estimado ¹ Modelo 1	D.Q.	Coef.estimado ¹ Modelo 4	D.Q.
25 - 29	2.0	0.0	2.0	0.6	1.0
30 - 34	4.2	2.2	1.0	1.4	1.9
35 - 39	7.1	4.8	0.7	6.3	0.1
40 - 44	15.1	8.3	3.1	13.3	0.2
45 - 49	28.6	16.3	5.3	23.8	0.8
50 - 54	42.8	14.3	19.0	42.9	0.0
55 - 59	58.6	28.7	15.3	44.7	3.3
60 - 64	83.0	51.1	12.3	86.5	0.1
65 - 69	137.6	60.0	43.8	150.2	1.2

1 : por 100 000 mulheres

2 : dados referentes ao Registro de Câncer de São Paulo (Anexo 13)

$$* : D.Q. = \frac{(O_i - E_i)^2}{E_i}$$

Verifica-se, através do Gráfico 3, que apresenta os resultados dos coeficientes de incidência para o câncer de estômago para o sexo feminino, que é a partir da faixa etária dos 50 a 54 anos que o Modelo 1 passa a ser inadequado e as diferenças entre os coeficientes observado e esperado passam a aumentar.

Já o Modelo 4 fornece uma curva muito próxima aos valores reais e está sempre oscilando entre eles.

GRÁFICO 3

COEFICIENTE DE INCIDENCIA DE CA DE ESTOMAGO, SEXO FEMININO, MUNICIPIO DE SAO PAULO, 1978.

Porém ao observar o Gráfico 4 (semi-log) constata-se que as diferenças são de menor magnitude do que as do sexo masculino, mostrando que o Modelo 1 foi melhor para estimar a incidência de câncer de estômago para o sexo feminino do que para o masculino, embora essas estimativas sempre sejam inferiores aos valores reais (em média, cerca de 50 % menores para ambos os sexos).

O Modelo 4 também obteve melhores estimativas para o sexo feminino do que para o masculino, porém em ambos os sexos foi mais adequado de que o Modelo 1.

As duas curvas estimadas são semelhantes aos dados reais e mostram uma tendência de incidência crescente com a idade.

GRÁFICO 4

COEFICIENTE DE INCIDENCIA DE CA DE ESTOMAGO, SEXO FEMININO, MUNICIPIO DE SAO PAULO, 1978.

A discussão feita para o sexo masculino sobre a influência das taxas de sobrevivência relativa nos dois modelos é válida também para o sexo feminino, o que explicaria os melhores resultados obtidos pelo Modelo 4.

As oscilações apresentadas pelos dois modelos podem ser explicadas pela oscilação existente no número de óbitos por câncer de estômago para ambos os sexos para o Município de São Paulo, em 1978. Porém, considera-se que no caso das estimativas dessa localização de câncer isso não foi importante e não prejudicou as estimativas.

A Tabela 3 apresenta os resultados das estimativas dos coeficientes de incidência da traquéia, brônquios e pulmão para o sexo masculino, fornecidos pelos Modelos 2 e 4.

Tabela 3 - Coeficientes de Incidência observado e estimados de câncer de traquéia, brônquios e pulmão no sexo masculino (por 100 000 homens), segundo idade, modelos e valores dos desvios quadráticos (D.Q.*) do teste de aderência pelo χ^2 . Município de São Paulo, 1978.

IDADE	Coeficiente ^{1,2} Observado	Coef.estimado ¹ Modelo 2	D.Q.	Coef.estimado ¹ Modelo 4	D.Q.
25 - 29	1.2	0.3	0.7	0.5	0.4
30 - 34	2.6	1.6	0.4	2.4	0.0
35 - 39	7.5	5.8	0.4	8.4	0.1
40 - 44	17.0	7.7	5.1	11.3	1.9
45 - 49	33.6	23.6	3.0	34.4	0.0
50 - 54	62.4	46.6	4.0	68.1	0.5
55 - 59	102.7	58.3	19.2	82.0	4.2
60 - 64	165.8	123.5	10.8	173.2	0.3
65 - 69	217.6	182.5	5.7	258.0	7.5

1 : por 100 000 homens

2 : dados referentes ao Registro de Câncer de São Paulo (Anexo 14)

$$* : D.Q. = \frac{(O_i - E_i)^2}{E_i}$$

Ao analisar os resultados do Modelo 2 nota-se que este modelo foi mais preciso para estimar a incidência de câncer do pulmão, para ambos os sexos do que o Modelo 1 o foi para o câncer do estômago. O χ^2 de aderência para o Modelo 2 foi de 49,3 ($p < 0,001$), ainda significativa, porém ao exame dos resultados pode ser verificado que as grandes diferenças foram nas faixas etárias de 50 a 54 anos e 55 a 59 anos, o que elevou o χ^2 final. Isso demonstra que os coeficientes de incidência para o câncer de pulmão fornecidos pelo Modelo 2, embora subestimados, são razoáveis para a maioria das faixas etárias.

O Modelo 4 fornece estimativas para a incidência de câncer de traquéia, brônquios e pulmão, para o sexo masculino, mais precisas do que o Modelo 2, com um χ^2 de aderência de 14,9 ($p < 0,10$), o que mostra o bom ajuste daquele modelo com os dados reais. Porém é interessante notar que também na faixa etária de 55 a 59 anos, à semelhança do Modelo 2, foi onde ocorreu uma grande diferença entre os valores estimados e os observados. Este fato propicia conjeturas a respeito da influência do número de óbitos na avaliação dos modelos.

Apesar de a incidência de câncer de traquéia, brônquios e pulmão ser sempre crescente com a idade, o número de óbitos nem sempre o é e isso se reflete nas estimativas feitas para os coeficientes de incidência que apresentam muitas oscilações nos dois modelos, sempre nas mesmas faixas etárias, que podem ser melhor visualizadas através dos gráficos. Devido ao menor número de óbitos para esta localização de câncer essas oscilações parecem influenciar mais a análise dos modelos.

Ao observar os Gráficos 5 e 6 onde aparecem as curvas dos coeficientes de incidência de câncer da traquéia, brônquios e pulmão para o sexo masculino em escala normal e semi-logarítmica, respectivamente, fica clara a proximidade das três curvas, com uma evidente maior precisão daquela obtida pelo Modelo 4.

GRÁFICO 5

COEFICIENTE DE INCIDENCIA DE CA DE PULMAO, SEXO MASCULINO, MUNICIPIO DE SAO PAULO, 1978.

É curioso notar que o formato das duas curvas estimadas no gráfico semi-log é igual, com uma diferença constante para todas as idades.

GRÁFICO 6

COEFICIENTE DE INCIDENCIA DE CA DE PULMAO, SEXO MASCULINO, MUNICIPIO DE SAO PAULO, 1978.

A diferença poderia ser explicada pelo fato de as taxas de sobrevivência utilizadas não serem adequadas para o Município de São Paulo, em 1978 o que faz com que haja um erro constante para todas as idades. Além disso outra hipótese pode ser suscitada que é a influência das taxas de sobrevivência relativa acumuladas que também são utilizadas pelo Modelo 2, pois à semelhança do Modelo 1, o Modelo 2 também acompanha uma coorte de pacientes sadios por vários anos.

Ao analisar-se os resultados das estimativas dos coeficientes de incidência do câncer da traquéia, brônquios e pulmão para o sexo feminino apresentados na Tabela 4, observa-se que os melhores resultados obtidos pelo Modelo 2 foram para esta localização e sexo, embora sub-enumerados. O teste de aderência aos valores reais para este modelo forneceu um χ^2 de 20,6 ($p < 0,002$). As grandes diferenças nas estimativas dos coeficientes de incidência se deram nas faixas etárias de 55 a 59 e de 65 a 69 anos.

Tabela 4 - Coeficientes de Incidência observado e estimados de câncer de traquéia, brônquios e pulmão no sexo feminino (por 100.000 mulheres), segundo idade, modelos e valores dos desvios quadráticos (D.Q.*) do teste de aderência pelo χ^2 . Município de São Paulo, 1978.

IDADE	Coeficiente ^{1,2} Observado	Coef.estimado ¹ Modelo 2	D.Q.	Coef.estimado ¹ Modelo 4	D.Q.
25 - 29	0.6	0.6	0.0	0.9	0.2
30 - 34	0.7	0.0	0.7	0.0	0.7
35 - 39	3.1	2.6	0.1	3.7	0.1
40 - 44	4.5	1.1	2.6	1.7	1.7
45 - 49	9.3	4.7	2.3	6.8	0.7
50 - 54	16.0	16.9	0.1	24.4	4.4
55 - 59	28.0	12.4	8.7	17.6	3.9
60 - 64	38.7	31.7	1.3	44.3	0.8
65 - 69	41.1	27.1	4.8	38.7	0.1

1 : por 100 000 mulheres

2 : dados referentes ao Registro de Câncer de São Paulo (Anexo 14)

$$(O_i - E_i)^2$$

$$* : D.Q. = \frac{(O_i - E_i)^2}{E_i}$$

Em que pese os bons resultados obtidos pelo Modelo 2, as estimativas dos coeficientes de incidência de câncer de traquéia, brônquios e pulmão para o sexo feminino fornecidas pelo Modelo 4 foram mais precisas. O χ^2 de aderência para esse último modelo foi de 12,6 ($p < 0,20$), mostrando a proximidade das estimativas aos valores reais. As maiores diferenças ocorreram nas faixas etárias de 50 a 54 anos e de 55 a 59 anos.

O Gráfico 7 representa os coeficientes de incidência para câncer de traquéia, brônquios e pulmão estimados para o sexo feminino. As oscilações que aparecem nas curvas referentes aos Modelos 2 e 4 estão nas mesmas faixas etárias e são devidos ao pequeno número de óbitos de câncer dessa localização para as mulheres. Apesar das estimativas serem próximas aos valores reais, os gráficos têm um formato irregular, principalmente nas idades mais avançadas. Nota-se que as estimativas fornecidas pelo Modelo 2 estão sempre sub-enumeradas e as do Modelo 4 oscilam entre os valores reais.

GRÁFICO 7

COEFICIENTE DE INCIDENCIA DE CA DE PULMAO, SEXO FEMININO, MUNICIPIO DE SAO PAULO, 1978.

As oscilações apresentadas pelas estimativas permanecem quando se observa o Gráfico 8, onde os coeficientes de incidência aparecem em escala logarítmica, principalmente na estimativa da faixa etária de 30 a 34 anos, que foi um valor zero para ambos os modelos, porém, ao se analisar as tendências, verifica-se que a incidência de câncer de traquéia, brônquios e pulmão para o sexo feminino é crescente com a idade nos dois modelos, acompanhando os valores reais.

GRÁFICO 8

COEFICIENTE DE INCIDENCIA DE CA DE PULMAO, SEXO FEMININO, MUNICIPIO DE SAO PAULO, 1978.

Aparentemente, o Modelo 2 sofre menos influência das taxas de sobrevivência relativa do que o Modelo 1, quando utilizado para estimar a incidência de câncer da traquéia, brônquios e pulmão, para ambos os sexos, embora sempre estejam sub-enumerados. A diferença entre o valor estimado e o esperado é muito variável, dependendo da faixa etária e sexo, mas, em sua maioria, gira em torno de 30 %.

Porém quando se observa que as curvas estimadas têm o mesmo formato, mas diferem por um valor constante em todas as faixas etárias, tem-se a impressão que a escolha das taxas de sobrevivência relativas não foi adequada e fez com que o Modelo 2 produzisse coeficientes de incidência subestimados.

Outro ponto importante a ser discutido e que pode estar influenciando mais intensamente no sentido negativo as estimativas do Modelo 2, talvez seja o pequeno número de óbitos e seu comportamento irregular com a idade.

O que faria com que o número de óbitos por câncer da traquéia, brônquios e pulmão não seja constantemente crescente com a idade, à semelhança da incidência? Esta questão poderia ser explicada através da presença de riscos competitivos, principalmente nas idades mais avançadas. Isso estaria influenciando a mortalidade por câncer e as pessoas estariam morrendo **com** o câncer e não **pelo** câncer. Esse efeito é mais sentido naquelas localizações em que o número de óbitos não é muito grande como é o caso do câncer de pulmão em mulheres no Município de São Paulo, em 1978.

As piores estimativas dos coeficientes de incidência realizadas tanto pelo Modelo 2 quanto pelo Modelo 4 foram para o câncer de mama feminina, como pode ser observado pelos resultados apresentados pela Tabela 5.

Tabela 5 - Coeficientes de Incidência observado e estimados de câncer de mama feminina (por 100.000 mulheres), segundo idade, modelos e valores dos desvios quadráticos (D.Q.*) do teste de aderência pelo χ^2 . Município de São Paulo, 1978.

IDADE	Coeficiente ^{1,2} Observado	Coef.estimado ¹ Modelo 2	D.Q.	Coef.estimado ¹ Modelo 4	D.Q.
25 - 29	8.2	0.4	7.4	3.0	3.3
30 - 34	28.4	18.7	3.3	23.4	0.9
35 - 39	53.3	26.5	13.5	48.6	0.4
40 - 44	90.1	23.7	48.9	65.7	6.6
45 - 49	145.7	63.1	46.8	125.5	2.8
50 - 54	185.3	25.5	137.8	126.1	18.9
55 - 59	178.4	99.1	35.2	183.3	0.1
60 - 64	203.6	53.9	110.1	193.1	0.5
65 - 69	227.0	114.5	55.8	265.4	6.5

1 : por 100 000 mulheres

2 : dados referentes ao Registro de Câncer de São Paulo (Anexo 15)

$$(O_i - E_i)^2$$

$$* : D.Q. = \frac{(O_i - E_i)^2}{E_i}$$

Para algumas das faixas etárias o coeficiente estimado pelo Modelo 2 foi duas vezes menor do que o real, porém para a faixa etária de 25 a 29 anos, ele foi vinte vezes menor e na de 50 a 54 anos foi quase 7 vezes menor. O χ^2 de aderência foi de 458,8 ($p < 0,001$), o maior de todas as estimativas feitas, mostrando o péssimo ajuste dos dados estimados aos reais.

O Modelo 4 continua fornecendo resultados mais precisos do que o Modelo 2, apesar de o χ^2 de aderência ter dado significativo: 40,0 ($p < 0,001$), . Interessante notar que as piores estimativas obtidas pelo modelo desenvolvido neste trabalho referem-se às faixas etárias dos 40 aos 54 anos, idade esta em que ocorrem intensas alterações hormonais nas mulheres. A partir da faixa etária dos 55 a 59 anos o Modelo 4 volta a estimar adequadamente os coeficientes de incidência de câncer de mama feminina.

Os melhores resultados apresentados pelo Modelo 4 em relação ao Modelo 2 poderiam ser explicados tanto pela escolha não adequada das taxas de sobrevivência relativa quanto pelo efeito acumulativo das mesmas, à semelhança do que ocorreu para as estimativas das outras localizações analisadas.

Porém o fato de que tanto o Modelo 2 quanto o Modelo 4 terem apresentado estimativas ruins para o câncer de mama permite tecer dois comentários: o primeiro é o de que a presença de riscos competitivos nas pacientes com esse tipo de câncer tenha uma influência mais acentuada nas idades mais avançadas e, conforme já sugerido anteriormente, as pacientes estariam morrendo **com** o câncer e não **pelo** câncer de mama.

Outro ponto é a dificuldade de se elaborar modelos matemáticos para algumas localizações de câncer que sofrem influência não só da idade, mas de outros fatores, principalmente endógenos. No caso do câncer de mama feminina é reconhecida a influência que os fatores reprodutivos e alterações hormonais exercem na sua incidência.

Isso demonstra a necessidade de se incorporar novos elementos ao Modelo 4, quando se desejar estimar incidência de câncer de localizações que não são função prioritariamente da idade. A solução seria, então, modificar o modelo proposto incorporando informações da teoria da carcinogênese, além de dados de população, óbito e sobrevivência.

Observa-se através do Gráfico 9 que há grandes oscilações das estimativas feitas pelo Modelo 2 para os coeficientes de incidência de câncer de mama feminina e, além disso, as diferenças entre as incidências observadas e estimadas acentuam-se fortemente com o avançar das idades. Essas diferenças não desaparecem mesmo quando as incidências estão em escala logarítmica (Gráfico 10), mostrando a inadequação deste modelo para estimar os coeficientes de incidência de câncer de mama feminina.

A curva obtida pelo Modelo 4, apesar de também apresentar oscilações, mostra-se muito próxima da curva com dados reais nas faixas etárias iniciais e a partir dos 55 anos. Há uma mudança no padrão da curva entre as idades de 40 a 54 anos, porém no restante é semelhante à dos dados reais.

GRÁFICO 9

COEFICIENTE DE INCIDENCIA DE CA DE MAMA FEMININA, MUNICIPIO DE SAO PAULO, 1978.

As oscilações podem ser explicadas pelo fato de o número de óbitos não ser crescente com a idade à semelhança da incidência de câncer de mama feminina.

GRÁFICO 10

COEFICIENTE DE INCIDENCIA DE CA DE MAMA FEMININA, MUNICIPIO DE SAO PAULO, 1978.

Várias conjecturas podem ser feitas para explicar o fato de os coeficientes de incidência estarem subestimados, tanto pelo Modelo 1 quanto pelo Modelo 2, conforme já comentado anteriormente.

Uma delas é a de que a escolha das taxas de sobrevivência não foi adequada. Sabe-se que elas dependem de fatores do tipo idade, sexo, raça, acesso e disponibilidade de recursos nos serviços de prevenção e tratamento de câncer. Além disso a definição do que é caso (ou seja, a partir de que estágio da doença o indivíduo é considerado caso) e a implantação de programas de *screening* (que podem apenas estar fornecendo um diagnóstico precoce, mas não um aumento na sobrevida pela doença) fazem com que a utilização de taxas de sobrevivência de outras localidades (geralmente mais desenvolvidas que o Município de São Paulo) não seja conveniente e prejudique a avaliação da adequação de modelos matemáticos para estimar a incidência de câncer.

Porém, dentre as poucas opções disponíveis, considera-se que as informações de sobrevivência fornecidas pelo Registro de Câncer de base populacional da Noruega, para o período de 1953 a 1967 foi a escolha mais adequada.

Uma outra hipótese poderia ser o efeito acumulativo das taxas de sobrevivência relativa. Esse efeito parece estar presente principalmente nas estimativas do Modelo 1.

Uma outra explicação para a falta de precisão de algumas estimativas poderia ser o comportamento estranho dos óbitos por câncer para as localizações escolhidas e o número pequeno dos mesmos, principalmente no caso de câncer da traquéia, brônquios e pulmão para o sexo feminino. Pode ser verificado que, apesar de os coeficientes de incidência reais serem crescentes com a idade para ambos os sexos, o mesmo não acontece com os coeficientes de mortalidade que apresentam oscilações não esperadas, principalmente no que se refere ao câncer de mama feminina.

Provavelmente não esteja ocorrendo grande número de erros na declaração da causa básica do óbito, fazendo com que o número de óbitos por essas localizações de câncer não esteja sub-enumerado. De acordo com PERCY, 1981, (53) em trabalho que verificou a acurácia das declarações de óbito nos Estados Unidos no período de 1970 e 1971, cuja causa básica era câncer de alguma localização, quando confrontadas com os registros hospitalares (inclusive relatórios de autópsia e de cirurgia) mostrou terem acima de 90% de precisão para as localizações estômago, traquéia, brônquios e pulmão e mama.

Logo a variabilidade do óbito poderia ser interpretada como sendo devida à presença de riscos competitivos de outras doenças ou de causas associadas ao câncer.

Como os modelos dependem dessa informação para o cálculo das estimativas, a sua avaliação fica, dessa maneira, prejudicada. Talvez isso poderia ser evitado se se utilizassem séries de mortalidade de períodos mais longos que um ano, mas não é o caso na avaliação desses modelos.

Não existe um modelo perfeito, como não existe um instrumento de medida perfeito, mas existe aquele que mais se aproxima da realidade e este parece ser o objetivo de quem trabalha na área de modelagem. As hipóteses levantadas nessa discussão devem ser testadas em futuros trabalhos.

VI. COMENTÁRIOS

1. O Modelo desenvolvido por Klementiev (Modelo 1) forneceu estimativas dos coeficientes de incidência de câncer de estômago próximos aos valores reais, tanto para o sexo masculino, quanto para o feminino até a idade de 49 anos, embora sempre sub-enumeradas. A partir dos 50 anos as diferenças se acentuaram e os coeficientes ficaram mais subestimados (em média os coeficientes reais são o dobro dos estimados). Essas diferenças mostraram-se menos acentuadas no sexo feminino. As curvas dos coeficientes de incidência estimados e real têm o mesmo formato, o que mostra que elas podem prever tendências. Isso faz levantar a hipótese de que a escolha das taxas de sobrevivência não tenha sido adequada e/ou o seu efeito cumulativo esteja interferindo nas estimativas, pois esse modelo acompanha uma coorte de indivíduos sadios durante vários anos até a sua morte.

2. O modelo desenvolvido por Verdecchia e colaboradores (Modelo 2), quando testado para o câncer da traquéia, brônquios e pulmão, mostrou um melhor ajuste com os coeficientes de incidência dessa localização para o Município de São Paulo, em 1978, para ambos os sexos. Porém também estão subestimados, principalmente nas faixas etárias de 55 a 59 anos e de 60 a 64 anos para o sexo masculino e de 55 a 59 anos para o sexo feminino. Contudo o comportamento das curvas dos coeficientes de incidência estimados e real é o mesmo e elas estão muito próximas, o que reforça a hipótese de não ser o modelo inadequado, mas, sim, a escolha das taxas de sobrevivência inconveniente e/ou o seu efeito cumulativo. Além disso, a irregularidade do comportamento do óbito com a idade pode estar interferindo na análise do modelo, principalmente em relação ao sexo feminino onde o número de óbitos é muito pequeno.

3. O modelo de Verdecchia e colaboradores não foi adequado para estimar os coeficientes de incidência do câncer de mama feminina para o Município de São Paulo, em 1978.

4. O modelo proposto por este trabalho (Modelo 4) foi o que forneceu as melhores estimativas dos coeficientes de incidência de estômago, traquéia, brônquios e pulmão e mama feminina para o Município de São Paulo, em 1978, quando comparado com os modelos de literatura. As piores estimativas fornecidas por este modelo referem-se ao câncer de mama feminina nas faixas etárias onde há uma alteração hormonal importante. Isso se explica pelo fato de essa localização de câncer ser difícil de ser modelada, pois depende de fatores reprodutivos e alterações hormonais, sugerindo que para se elaborar modelos matemáticos para estimar a incidência de câncer de mama devem ser incorporadas informações da teoria da carcinogênese.

5. Ao utilizar-se dados de mortalidade, população e sobrevivência de diferentes fontes, não necessariamente compatíveis, para se estimar coeficientes de incidência que são fornecidos por uma quarta fonte, pode-se chegar a resultados não satisfatórios devido à inadequação da escolha das fontes. Porém, às vezes, não há opção para escolha, tendo-se que lançar mão dos dados existentes.

6. Esta é uma importante linha de pesquisa a ser desenvolvida e a continuação dos estudos nessa área deve envolver os seguintes aspectos:

6.1. Procurar interação entre os diversos modelos de literatura e o modelo desenvolvido por este trabalho em um único modelo mais completo e próximo da realidade.

6.2. Criar um novo modelo, dentro da teoria de processos estocásticos, utilizando dados de população, óbito e sobrevivência, associados à teoria da carcinogênese. Esse modelo deve contar com mais de um estágio de transição, considerar os diversos estágios da carcinogênese (indução e promoção), ter estágios que possam ser considerados reversíveis, considerar os diversos estadiamentos da doença, envolver maior complexidade na competição entre os vários riscos de morte na vida de um indivíduo (isto é, a probabilidade de morrer por uma doença é influenciada pela competição de outros riscos de morte) e estudar o efeito das causas eventualmente associadas.

6.3. Calcular ou estimar taxas de sobrevivência de câncer para o Município de São Paulo, pois elas dependem de fatores do tipo idade, sexo, raça, acesso e disponibilidade de recursos nos serviços de prevenção e tratamento de câncer. Além disso, a definição de caso e a implantação de programas de *screening* fazem com que a utilização de taxas de sobrevivência de outras localidades (geralmente mais desenvolvidas que o Município de São Paulo) não seja conveniente e prejudique a avaliação da adequação de modelos matemáticos para estimar a incidência de câncer.

VII. REFERÊNCIAS BIBLIOGRÁFICAS

1. ANDERSON,D.H. & ROLLER,T. Equilibrium points for nonlinear compartmental models. *Math.Biosci.*, **103**:159-201, 1991.
2. ANUÁRIO ESTATÍSTICO DO ESTADO DE SÃO PAULO - 1988. São Paulo, Fundação SEADE, 1989.
3. ARMITAGE,P. & DOLL, R. The age distribution of cancer and multi-stage theory of carcinogenesis. *Brit. J.Cancer*, **8**:1-12, 1954.
4. ARMITAGE,P. & DOLL, R. A two-stage theory of carcinogenesis in relation to the age distribution of human cancer. *Brit. J. Cancer*, **11**:161-169, 1957.
5. BENM,H. Determinantes económicos y sociales de la mortalidad en America Latina. *Rev. cubana Adm. publ.*, **6**: 1-44, 1980.
6. BOSCHI-PINTO, C. Mortalidade por câncer de trato digestivo no Estado do Rio de Janeiro: uma apreciação por estratificação. Rio de Janeiro, 1991./Dissertação de Mestrado - Escola Nacional de Saúde Pública- FIOCRUZ/
7. BOSS,L.P. & SUAREZ,L. Uses of data to plan cancer prevention and control programs. *Publ. Hlth Rep.*, **105**: 354-360, 1990.
8. BRENNAN, M.E. & CLARE, P.H. The relationship between mortality and two indicators of morbidity. *J.Epidem.Community Hlth*, **34** : 134-138, 1980.
9. BROWN,C.C. & CHU,K.C. Additive and multiplicative models and multistage carcinogenesis theory. *Risk Anal.*, **9**:99-105, 1989.
10. CALDERÔN, C.P. & KWEMBE, T.A. Modeling tumor growth. *Math. Biosci.*, **103**:97-114, 1991.
11. CHIANG, C.L. *Introduction to stochastic processes in biostatistics*. New York, John Wiley & Sons, 1968.
12. CHIANG, C.L. & COHEN, R.D. How to measure health: a stochastic model for an index of health. *Int. J. Epidem.*, **2**: 7-13, 1973.
13. CHIANG, Y.-K. et al. An illness-death process with time-dependent covariates. *Biometrics*, **45**:669-681, 1989.
14. CHU, K.C. et al. Estimating cancer mortality rates from SEER incidence and survival data. *Publ.Hlth Rep.*, **105**:36-46, 1990.
15. CIAMPI, A. et al. Regression and recursive partition strategies in the analysis of medical survival data. *J.clin.Epidem.*, **41**:737-748, 1988.
16. CLASSIFICAÇÃO Internacional de Doenças; 9ª revisão, 1975. São Paulo, Centro da OMS para Classificação de Doenças em Português, 1978/79. 2v.

17. COCHRANE, A.S. et al. Health service "input" and mortality "output" in developed countries. *J. Epidem. Community Hlth*, **32**: 200-205, 1978.
18. COOK, P.J. et al. A mathematical model for the age distribution of cancer in man. *Int.J.Cancer*, **4**:93-112, 1969.
19. DAMIANI, P. Études sur la morbidité. *Rev. Statist. appl.*, **21**: 33-58, 1973.
20. DAMIANI, P. Mesure de la morbidité: liaison avec la mortalité. *Int. Statist. Rev.*, **45**: 39-50, 1977.
21. DAMIANI, P. et al. Évaluation de la morbidité a partir de la mortalité. *J. Soc. Statist. Paris*, **123**: 58-74, 1982.
22. DEWANJI, A. et al. A stochastic two-stage model for cancer risk assessment. II. The number and size of premalignant clones. *Risk Anal.*, **9**: 179-187, 1989.
23. ELANDT-JOHNSON, R.C. & JOHNSON, N.L. *Survival models and data analysis*. New York, John Wiley & Sons, 1980.
24. FEINSTEIN, A.R. et al. Mathematical models and scientific reality in occurrence rates for diseases. *Amer. J. publ. Hlth*, **79**: 1303-1304, 1989.
25. FIX, E. & NEYMAN, J. A simple stochastic model of recovery, relapse, death and loss of patients. *Hum. Biol.*, **23**: 205-241, 1951.
26. FOSTER, D.P. Mortality, morbidity and resource allocation. *Lancet*, **1**:997-998, 1977.
27. GANDY, R.J. A comparison of two proxy measures for morbidity. *Epidem. Community Hlth*, **33**: 100-103, 1979.
28. GRAVELLE, H.S.E. & BACKHOUSE, M.E. International cross-section analysis of the determination of mortality. *Soc.Sci.Med.*, **25**: 427-441, 1987.
29. GREENLAND, S. Commentary: modeling and variable selection in epidemiologic analysis. *Amer. J. publ. Hlth*, **79**: 340-349, 1989.
30. GROSS, A.J. et al. Estimation of survival parameters when one of two organs must function for survival. *Biometrics*, **27**: 369-377, 1971.
31. GRUNDY, F. & REINKE, W.A. *Procedimentos de pesquisa em saúde e métodos matemáticos de gestão*. Genebra, Organização Mundial de Saúde, 1974.
32. FINNISH CANCER REGISTRY. Survival of cancer patients in Finland in 1953-1974. *Ann.clin. Res.*, Helsinki, **13** (Suppl.31), 1981.
33. HAKULINEN, T. A comparison of nationwide cancer survival statistics in Finland and Norway. *Wld Hlth Statist. Quart.*, **36**: 35-46, 1983.
34. KAMPERT, J.B. & SILVERS, A. Goodness of fit in Armitage-Doll Poisson Models. *Biometrics*, **45**:1271-1280, 1989.
35. KATZOFF, M. The application of time series forecasting methods to an estimations problem using joint surveillance statistics. *Statist. Med.*, **8**: 333-341, 1989.

36. LAURENTI, R. et al. *Estatísticas de saúde*. 2a. ed. São Paulo, EPU, 1987.
37. MANTON, K.G. et al. Life table techniques for multiple-cause mortality. *Demography*, **13**: 541-564, 1976.
38. MANTON, K.G. & STALLARD, E. Maximum Likelihood estimation of a stochastic compartment model of cancer latency: lung cancer mortality among white females in U.S. *Comput. biomed. Res.*, **12**: 313-325, 1979.
39. MANTON, K.G. & STALLARD, E. The use of mortality time series data to produce hypothetical morbidity distributions and project mortality trends. *Demography*, **19**: 223-240, 1982.
40. MASSAD, E. Sobre a biomatemática das moléstias infecciosas: um estudo teórico exemplificado por dois modelos originais. São Paulo, 1987. /Tese de Livre-Docência - Faculdade de Medicina da USP/
41. MAZUMDAR, S. et al. Multistage modeling of lung cancer mortality among arsenic-exposed copper-smelter workers. *Risk Anal.*, **9**: 551-563, 1989.
42. MEYER, P.L. *Probabilidade: aplicações à estatística*. Rio de Janeiro, Ao Livro Técnico/ Ed. USP 1969.
43. MINISTÉRIO DA SAÚDE. *Estatísticas de mortalidade. Brasil - 1985*. Brasília, 1988.
44. MIRRA, A.P. & FRANCO, E.L. *Cancer incidence in São Paulo County - Brazil - (1969-1973-1978)*. São Paulo, Registro de Câncer de São Paulo/Instituto Ludwig de Pesquisa sobre o Câncer, 1985.
45. MIRRA, A.P. & FRANCO, E.L. *Cancer mortality in São Paulo - Brazil*. São Paulo, Registro de Câncer de São Paulo/Instituto Ludwig de Pesquisa sobre o Câncer, 1987. (LICR Cancer Epidemiology Monograph Series - v.3).
46. MOOLGAVKAR, S.H. & VENZON, D.J. Two-event models for carcinogenesis: incidence curves for childhood and adult tumors. *Math. Biosci.*, **47**: 55-77, 1979.
47. MOOLGAVKAR, S.H. & KNUDSON JR, A.G. Mutation and cancer: a model for human carcinogenesis. *J. Nat. Cancer Inst.*, **66**: 1037-1052, 1981.
48. MOOLGAVKAR, S.H. Carcinogenesis modeling: from molecular biology to epidemiology. *Ann. Rev. publ. Hlth*, **7**: 151-169, 1986.
49. MOOLGAVKAR, S.H. et al. A stochastic two-stage model for cancer risk assessment. I. The hazard function and the probability of tumor. *Risk Anal.*, **8**: 383-392, 1988.
50. NATIONAL CENTER FOR HEALTH STATISTICS. *An index of health: mathematical models*. Washington, D.C., 1965. (Public Health Service Publication nº 1000, Series 2, nº 5).
51. NOONEY, G.C. Mathematical models, reality and results. *J. theoret. Biol.*, **9**: 239-252, 1965.

52. NORWEGIAN CANCER SOCIETY. *The Cancer Registry of Norway: survival of cancer patients: cases diagnosed in Norway: 1953-1967.* Oslo, 1975.
53. PERCY, C. et al. Accuracy of cancer death certificates and its effect on cancer mortality statistics. *Amer.J.publ.Hlth*, 71:242-250, 1981.
54. PUFFER, R.R. New approaches for epidemiologic studies of mortality statistics. *Bull. Pan Amer. Hlth. Org.*, 23: 365-383, 1989.
55. ROBINS, J. & GREENLAND, S. The probability of causation under a stochastic model for individual risk. *Biometrics*, 45: 1125-1138, 1989.
56. RUZICKA, L.T. The use of mortality and morbidity statistics for national health promotion. Geneva, World Health Organization, 1990. (DES/ICD/C/90.17) / Apresentado no Meeting of Heads of WHO Collaborating Centres for The Classification of Diseases, London, 1990./
57. SAAD, P.M. Mortalidade infantil por causas no Estado de São Paulo (Brasil) em 1983. *Rev.Saúde públ.*, S.Paulo, 20:481-488, 1986.
58. SAAD, P.M. Mortalidade por causas múltiplas. *S. Paulo Perspect.*, 1(3): 77-83, 1987.
59. SANTO, A.H. Causas múltiplas de morte: formas de apresentação e métodos de análise. São Paulo, 1988. /Tese de Doutorado - Faculdade de Saúde Pública da USP/.
60. SILCOCKS, P.B.S. et al. Making cancer statistics more informative: measures of the quality of recorded diagnosis in a population-based registry. *Eur.J.Cancer clin.Oncol.*, 25: 1467-1473, 1989.
61. SWAROOPS, S. *Estadística sanitaria.* México, Fondo de Cultura, 1964.
62. THORSLUND, T.W. et al. Biologically motivated cancer risk models. *Risk Anal.*, 7: 109-119, 1987.
63. UNIÃO INTERNACIONAL CONTRA O CANCER (UICC). *Manual de oncologia clínica.* 5ª ed. São Paulo, Fundação Oncocentro de São Paulo, 1991.
64. VERDECCHIA, A. et al. Estimation of cancer morbidity using mortality data. *Tumori*, 71: 431-439, 1985.
65. VERDECCHIA, A. et al. A method of the estimation of chronic disease morbidity and trends from mortality data. *Statist. Med.*, 8: 201-216, 1989.

ANEXOS

ANEXO 1

Estimativa da população do Município de São Paulo, segundo sexo e faixa etária, 1978.

FAIXA ETÁRIA	SEXO	
	MASCULINO	FEMININO
0 - 4	385 681	374 604
5 - 9	324 747	318 791
10 - 14	295 949	299 965
15 - 19	325 989	354 745
20 - 24	388 831	394 749
25 - 29	346 510	355 794
30 - 34	275 420	286 620
35 - 39	213 945	225 642
40 - 44	188 540	199 117
45 - 49	157 782	171 948
50 - 54	140 130	157 301
55 - 59	103 713	118 340
60 - 64	73 845	90 769
65 - 69	55 337	70 806
70 - 74	45 048	66 447
75 -	23 403	37 686
TOTAL	3 344 870	3 523 324

FONTE:44

ANEXO 2

Número de óbitos segundo idade e sexo de residentes no Município de São Paulo, 1978

FAIXA ETÁRIA	SEXO	
	MASCULINO	FEMININO
0 - 4	8 943	6 978
5 - 9	223	155
10 - 14	282	164
15 - 19	639	290
20 - 24	909	375
25 - 29	887	452
30 - 34	910	518
35 - 39	1 076	552
40 - 44	1 364	683
45 - 49	1 740	951
50 - 54	2 102	1 127
55 - 59	2 090	1 238
60 - 64	2 202	1 614
65 - 69	2 433	1 806
70 - 74	2 359	2 036
75 -	4 023	5 128
TOTAL	32 182	24 067

FONTE: Fundação SEADE - dados brutos

ANEXO 3

Coeficiente de mortalidade (por 10 000) segundo idade e sexo. Município de São Paulo, 1978

FAIXA ETÁRIA	SEXO	
	MASCULINO ¹	FEMININO ²
25 - 29	255.98	127.04
30 - 34	330.40	180.73
35 - 39	502.92	244.64
40 - 44	723.45	343.01
45 - 49	1 102.79	553.07
50 - 54	1 500.04	716.46
55 - 59	2 015.18	1 046.14
60 - 64	2 981.92	1 778.14
65 - 69	4 396.70	2 550.63
70 - 74	5 236.64	3 064.10

1: por dez mil homens

2. por dez mil mulheres

ANEXO 4

Coeficiente específico de mortalidade por câncer de estômago (por 100.000), segundo sexo e faixa etária. Município de São Paulo, 1978

FAIXA ETÁRIA	SEXO	
	MASCULINO ¹	FEMININO ²
0 - 4	0.00	0.00
5 - 9	0.00	0.00
10 - 14	0.00	0.00
15 - 19	0.00	0.00
20 - 24	0.26	0.76
25 - 29	0.00	0.28
30 - 34	1.45	0.70
35 - 39	8.42	3.10
40 - 44	16.45	6.52
45 - 49	33.60	11.62
50 - 54	50.69	20.96
55 - 59	79.09	22.79
60 - 64	101.60	44.02
65 - 69	157.28	76.18
70 - 74	182.10	72.16
75 -	457.38	209.39

FONTE: 45

1 : por cem mil homens

2 : por cem mil mulheres

ANEXO 5

Coeficiente específico de mortalidade por câncer de traquéia, brônquios e pulmão (por 100 000), segundo sexo e faixa etária. Município de São Paulo, 1978

FAIXA ETÁRIA	SEXO	
	MASCULINO ¹	FEMININO ²
0 - 4	0.00	0.00
5 - 9	0.00	0.00
10 - 14	0.00	0.00
15 - 19	0.00	0.00
20 - 24	0.51	0.00
25 - 29	0.29	0.56
30 - 34	1.45	0.00
35 - 39	5.14	2.21
40 - 44	6.89	1.00
45 - 49	20.89	4.07
50 - 54	41.34	14.61
55 - 59	52.00	10.97
60 - 64	109.55	27.51
65 - 69	162.44	23.98
70 - 74	161.44	33.07
75 -	315.80	66.26

FONTE: 45

1 : por cem mil homens

2 : por cem mil mulheres

ANEXO 6

Coeficiente específico de mortalidade por câncer de mama feminina (por 100 000 mulheres), segundo faixa etária. Município de São Paulo, 1978

FAIXA ETÁRIA	COEFICIENTE (POR 100.000 MULHERES)
0 - 4	0.00
5 - 9	0.00
10 - 14	0.00
15 - 19	0.00
20 - 24	0.51
25 - 29	0.84
30 - 34	6.62
35 - 39	13.72
40 - 44	18.56
45 - 49	35.44
50 - 54	35.56
55 - 59	60.77
60 - 64	63.83
65 - 69	87.47
70 - 74	55.62
75 -	153.73

FONTE: 45

ANEXO 7

Número de óbitos por câncer de estômago , segundo faixa etária e sexo de residentes no Município de São Paulo, 1978

FAIXA ETÁRIA	SEXO	
	MASCULINO	FEMININO
0 - 4	-	-
5 - 9	-	-
10 - 14	-	-
15 - 19	-	-
20 - 24	1	3
25 - 29	-	1
30 - 34	4	2
35 - 39	18	7
40 - 44	31	13
45 - 49	53	20
50 - 54	71	33
55 - 59	82	27
60 - 64	75	40
65 - 69	87	54
70 - 74	82	48
75 -	107	79
TOTAL	611	327

FONTE : 45

ANEXO 8

Número de óbitos por câncer da traquéia, brônquios e pulmão, segundo faixa etária e sexo de residentes no Município de São Paulo, 1978

FAIXA ETÁRIA	SEXO	
	MASCULINO	FEMININO
0 - 4	-	-
5 - 9	-	-
10 - 14	-	-
15 - 19	-	-
20 - 24	2	-
25 - 29	1	2
30 - 34	4	-
35 - 39	11	5
40 - 44	13	2
45 - 49	33	7
50 - 54	58	23
55 - 59	54	13
60 - 64	81	25
65 - 69	90	17
70 - 74	73	22
75 -	74	25
TOTAL	494	141

FONTE: 45

ANEXO 9

Número de óbitos por câncer de mama, segundo faixa etária de mulheres residentes no Município de São Paulo, 1978

FAIXA ETÁRIA	NÚMERO DE ÓBITOS
0 - 4	-
5 - 9	-
10 - 14	-
15 - 19	-
20 - 24	2
25 - 29	3
30 - 34	19
35 - 39	31
40 - 44	37
45 - 49	61
50 - 54	56
55 - 59	72
60 - 64	58
65 - 69	62
70 - 74	37
75 -	58
TOTAL	496

FONTE: 45

ANEXO 10

Taxa de sobrevivência relativa de câncer de estômago, segundo idade e sexo, de casos registrados no Registro de Câncer da Noruega, no período de 1953 a 1967

FAIXA ETÁRIA	TAXA SOBR.RELATIVA PARA 5 ANOS		TAXA SOBR.RELATIVA PARA 1 ANO	
	MASCULINO	FEMININO	MASCULINO	FEMININO
20 - 24	0.14	0.10	0.39	0.35
25 - 29	0.14	0.10	0.39	0.35
30 - 34	0.14	0.10	0.39	0.35
35 - 39	0.14	0.10	0.39	0.35
40 - 44	0.14	0.10	0.39	0.35
45 - 49	0.14	0.10	0.39	0.35
50 - 54	0.14	0.10	0.39	0.35
55 - 59	0.11	0.12	0.30	0.31
60 - 64	0.11	0.12	0.30	0.31
65 - 69	0.11	0.12	0.30	0.31
70 - 74	0.11	0.12	0.30	0.31

FONTE: 52

ANEXO 11

Taxa de sobrevivência relativa de câncer da traquéia, brônquios e pulmão, segundo idade e sexo, de casos registrados no Registro de Câncer da Noruega, no período de 1953 a 1967

FAIXA ETÁRIA	TAXA SOBR.RELATIVA PARA 5 ANOS		TAXA SOBR.RELATIVA PARA 1 ANO	
	MASCULINO	FEMININO	MASCULINO	FEMININO
20 - 24	0.12	0.14	0.28	0.32
25 - 29	0.12	0.14	0.28	0.32
30 - 34	0.12	0.14	0.28	0.32
35 - 39	0.12	0.14	0.28	0.32
40 - 44	0.12	0.14	0.28	0.32
45 - 49	0.12	0.14	0.28	0.32
50 - 54	0.12	0.14	0.28	0.32
55 - 59	0.06	0.09	0.25	0.25
60 - 64	0.06	0.09	0.25	0.25
65 - 69	0.06	0.09	0.25	0.25
70 - 74	0.06	0.09	0.25	0.25

FONTE: 52

ANEXO 12

Taxa de sobrevivência relativa de câncer de mama feminina, segundo idade, de casos registrados no Registro de Câncer da Noruega, no período de 1953 a 1967

FAIXA ETÁRIA	TAXA DE SOBREVIVÊNCIA	
	para 5 anos	para 1 ano
20 - 24	0.67	0.92
25 - 29	0.67	0.92
30 - 34	0.67	0.92
35 - 39	0.67	0.92
40 - 44	0.67	0.92
45 - 49	0.67	0.92
50 - 54	0.67	0.92
55 - 59	0.60	0.88
60 - 64	0.60	0.88
65 - 69	0.60	0.88
70 - 74	0.60	0.88

FONTE: 52

* : categoria 174 - CID 9ª revisão

ANEXO 13

Coeficiente de incidência (por 100 000) de câncer de estômago segundo faixa etária e sexo, de casos registrados no Registro de Câncer de São Paulo, 1978

FAIXA ETÁRIA	SEXO	
	MASCULINO ¹	FEMININO ²
0 - 4	0.0	0.0
5 - 9	0.0	0.0
10 - 14	0.0	0.0
15 - 19	0.0	0.0
20 - 24	0.3	1.0
25 - 29	1.2	2.0
30 - 34	2.6	4.2
35 - 39	12.2	7.1
40 - 44	34.5	15.1
45 - 49	50.1	28.6
50 - 54	107.5	42.8
55 - 59	139.5	58.6
60 - 64	191.6	83.0
65 - 69	310.0	137.6
70 - 74	378.3	190.2
75 -	433.9	207.4
TOTAL (AJUSTADO)	53.6	25.1

FONTE: 44

1: por cem mil homens

2. por cem mil mulheres

ANEXO 14

Coeficiente de incidência (por 100 000) de câncer da traquéia, brônquios e pulmão, segundo faixa etária e sexo, de casos registrados no Registro de Câncer de São Paulo, 1978

FAIXA ETÁRIA	SEXO	
	MASCULINO ¹	FEMININO ²
0 - 4	0.0	0.0
5 - 9	0.0	0.0
10 - 14	0.0	0.0
15 - 19	0.6	0.0
20 - 24	1.0	0.8
25 - 29	1.2	0.6
30 - 34	2.6	0.7
35 - 39	7.5	3.1
40 - 44	17.0	4.5
45 - 49	33.6	9.3
50 - 54	62.4	16.0
55 - 59	102.7	28.0
60 - 64	165.8	38.7
65 - 69	217.6	41.1
70 - 74	258.9	42.0
75 -	285.2	50.4
TOTAL (AJUSTADO)	36.5	7.9

FONTE: 44

1: por cem mil homens

2. por cem mil mulheres

ANEXO 15

Coeficiente de incidência (por 100 000 mulheres) de
câncer de mama, segundo faixa etária, de casos
registrados no Registro de Câncer de São Paulo, 1978

FAIXA ETÁRIA	Coefic. de Incid. (por 10 ⁵ mulheres)
0 - 4	0.0
5 - 9	0.0
10 - 14	0.0
15 - 19	0.3
20 - 24	2.0
25 - 29	8.2
30 - 34	28.4
35 - 39	53.3
40 - 44	90.1
45 - 49	145.7
50 - 54	185.3
55 - 59	178.4
60 - 64	203.6
65 - 69	227.0
70 - 74	246.2
75 -	274.6
TOTAL(AJUSTADO)	65.5

FONTE: 44