**UNIVERSIDADE DE SÃO PAULO
FACULDADE DE MEDICINA DE RIBEIRÃO PRETO**

MÁRCIO HENRIQUE CARVALHO GRADE

**“Avaliação da concordância em diferentes fases do processo diagnóstico,
embasada na necropsia, no HCFMRPUSP, nos anos de 1980, 1990 e 2000”**

RIBEIRÃO PRETO

2007

MÁRCIO HENRIQUE CARVALHO GRADE

**“Avaliação da concordância em diferentes fases do processo diagnóstico,
embasada na necropsia, no HCFMRPUSP, nos anos de 1980, 1990 e 2000”**

Tese apresentada ao Departamento de Patologia da
Faculdade de Medicina de Ribeirão Preto - USP, para
obtenção do Título de Doutor em Patologia.

Área de concentração: Patologia Experimental.

Orientador: Prof. Dr. Sérgio Britto Garcia

RIBEIRÃO PRETO

2007

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

FICHA CATALOGRÁFICA

Grade, Márcio Henrique Carvalho

Avaliação da concordância em diferentes fases do processo diagnóstico, embasada na necropsia, no HCFMRPUSP, nos anos de 1980, 1990 e 2000. Ribeirão Preto, 2007.

119 p.; 30cm

Tese de Doutorado, apresentada à Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo. Área de concentração: Patologia Experimental.

Orientador: Garcia, Sérgio Britto.

1. Processo diagnóstico. 2. Discrepâncias diagnósticas. 3. Necropsia.
4. Falibilidade. 5. Erro.

FOLHA DE APROVAÇÃO

Márcio Henrique Carvalho Grade

Avaliação da concordância em diferentes fases do processo diagnóstico, embasada na necropsia, no HCFMRPUSP, nos anos de 1980, 1990 e 2000.

Tese apresentada ao Departamento de Patologia da Faculdade de Medicina de Ribeirão Preto - USP, para obtenção do Título de Doutor.

Área de concentração: Patologia Experimental.

Aprovado em:

Banca Examinadora

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Agradecimentos

Ao Prof. Dr. Sérgio Britto Garcia por sua confiança,
amizade, pela oportunidade e por partilhar conceitos científicos e de vida.

À Prof^a Dr^a Carmen Cinira Santos Martim por sua amizade
e por propiciar meu desenvolvimento profissional com seu convívio.

Ao Departamento de Patologia,
seus funcionários e pós-graduandos por oferecer sua estrutura e apoio.

À Prof^a Dr^a Fabíola Reis de Oliveira
por sua participação na revisão dos prontuários e do texto,
por sua participação na minha vida e pelo presente.

A meus pais, meus irmãos,
minha família em Londrina e Ribeirão Preto e aos meus amigos.

Resumo

GRADE, M. H. C. **Avaliação da concordância em diferentes fases do processo diagnóstico, embasada na necropsia, no HCFMRPUSP, nos anos de 1980, 1990 e 2000.** 2007, 119p. Tese de Doutorado - Departamento de Patologia da Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2007.

O processo diagnóstico na medicina traduz-se por uma complexa interação entre conhecimentos, habilidades e procedimentos técnicos em condições de incerteza. As variáveis que o cercam são inúmeras e de difícil caracterização, dificultando o desenvolvimento de estudos mais aprofundados sobre o processo de elaboração do diagnóstico. O papel da necropsia na avaliação da performance do diagnóstico clínico já está bem demonstrado, sendo geralmente aceito que ela pode ser usada como uma ferramenta para analisar possíveis discrepâncias. A literatura médica vem apresentando com maior frequência a questão dos erros no cuidado à saúde, e têm proposto que novos estudos são necessários para avaliar os erros diagnósticos e com isto propiciar condições para uma melhora desta questão. Nós propusemos um estudo para verificar a concordância diagnóstica em três fases deste processo, em pacientes do Hospital das Clínicas de Ribeirão Preto. Analisamos retrospectivamente todos os diagnósticos de trezentos pacientes que faleceram em nosso hospital, sendo cem pacientes para cada ano escolhido para este estudo (1980, 1990 e 2000). Classificamos os diagnósticos em: clínicos iniciais, considerados aqueles de entrada do paciente sem o apoio de exames complementares dentro do hospital; clínicos finais, todos os realizados durante a internação e após o óbito; necroscópicos, todos os diagnósticos elencados no relatório final deste procedimento. Utilizamos três grandes grupos para comparar estas fases: concordância, discordâncias maiores e discordâncias menores. Com estes achados realizamos uma avaliação do processo diagnóstico. As doenças do aparelho circulatório formaram o maior grupo de causas básicas de óbito seguido pelas neoplasias. O percentual de realização de necropsias foi ao redor de 75%. O número total de procedimentos diagnósticos pesquisados aumentou de 236 em 1980 para 366 em 2000, principalmente devido a exames não invasivos como as ultrassonografias e as tomografias. As discrepâncias maiores não apresentaram melhora percentual, mantendo-se estáveis ao redor de 21% no período do estudo. Houve um aumento das concordâncias diagnósticas (de 41 para 55%) confrontando o diagnóstico clínico final e o necroscópico, porém às custas de diminuição das discrepâncias menores. Das variáveis analisadas de idade, gênero, tempo de estadia hospitalar, admissões anteriores, vulnerabilidade, exames comuns, modernos e totais, apenas a de admissões apresentou diferença estatística na comparação de pacientes com discrepâncias maiores em relação aos demais, com $p=0,006$, tendo o grupo dos demais pacientes uma média maior (de 1,85 x 1,35). A avaliação do processo diagnóstico através da comparação entre o diagnóstico clínico inicial e o clínico final mostrou concordância entre os passos em cerca de 80%. Os casos com discrepâncias maiores à necropsia em relação ao diagnóstico clínico final, apresentaram 90% de concordância entre o diagnóstico clínico inicial e o final. Este achado reforça a tese de que o encerramento prematuro do diagnóstico parece ser o principal fator relacionado com a manutenção destas discrepâncias. O aumento de concordância diagnóstica deu-se, portanto, com a diminuição de discrepâncias menores, sem alteração nas maiores que são aquelas que poderiam ter influenciado no prognóstico do paciente. A disponibilidade de novos exames e tecnologias de ponta não foi suficiente para alterar este quadro.

Abstract

GRADE, M. H. C. **Assessment of the agreement on different phases of the diagnostic process, based on necropsy, at HCFMRPUSP, in 1980, 1990 and 2000.** 2007, 119p. Tese de Doutorado - Departamento de Patologia da Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2007.

The diagnostic process in medicine is a complex interaction of cognitive skills and technical procedures in conditions of uncertainty. The variables around of it are innumerable and with difficult characterization. Because of this, there is not a good development of studies concerning the elaboration of diagnoses. The role of autopsy in auditing clinical diagnosis performance it is well demonstrated and it is generally accepted that autopsy may be used as a tool for quality management to analyze diagnostic discrepancies. The medical literature presents more frequently the question of medical errors. It has been proposed that studies are needed to assess how and when diagnostic errors occur and this might improve diagnostic performance further. We propose a study to verify the agreement of medical diagnostics, based on detailed, ongoing comparison of three steps of the clinical diagnoses with autopsy diagnoses over 20 years, in patients of the Ribeirão Preto Clinics Hospital. We analyzed retrospectively diagnostic errors, with use of necropsy as the gold standard for diagnosis. We randomly selected 300 patients who died at a tertiary-care teaching hospital in Brazil—100 in each of 1980, 1990, and 2000. We classified discrepancies between admission diagnosis, clinical diagnosis and necropsy findings, into three categories, major and minor discrepancies and concordance. With these findings we made an evaluation of the diagnostic process. The cardiovascular diseases were the major group of basic causes of death, and the neoplasias were the second. The percentage of necropsy was around 75%. The total number of diagnostics procedures increase from 236 in 1980 to 366 in 2000, mainly because of non invasive exams such ultrasonography and tomography. The major discrepancies maintained 21% over the studied period. There was an increase of the diagnostic concordance over the studied period (from 41 to 55%), when we compare the final clinical diagnostics and the necropsy, however only with the reduction of minor discrepancies. We analyzed variables like age, sex, days of stay on hospital, previous admissions, vulnerability, common exams, modern exams and total exams. Only previous admissions presented significant statistically difference, when comparing patients with major discrepancies with the others, $p=0,006$. The others group presented a higher media (1,85 x 1,35) of previous admissions. The assessment of the diagnostic process with the comparison of admission and the final clinical diagnoses revealed 80% of concordance. The cases of major discrepancies presented 90% of concordance between the admission and the final clinical diagnoses. This reinforces the thesis that the premature closure of the diagnostic is the principal factor related with the maintenance of these discrepancies. The increase of diagnostics concordances was due to the decrease of minor discrepancies, without alterations of majors. These ones are that with some influence in the prognostic of the patients. The availability of new exams and technologies were not sufficient to change the rate of discrepancies.

Sumário

INTRODUÇÃO	8
A falibilidade do processo diagnóstico	9
A necropsia e as discrepâncias diagnósticas	14
OBJETIVOS	25
MATERIAL	27
Local do estudo	28
Pacientes	29
Fontes de dados	29
Comitê de Ética	30
MÉTODOS.....	31
Amostragem	32
Coleta dos dados	32
Critérios de exclusão	34
Categorias comparativas	34
Classificação	35
Análise estatística.....	39
RESULTADOS	43
DISCUSSÃO	59
Caracterização da amostra	60
Causas básicas dos óbitos	61
Exame necroscópico completo.....	63
Comparação das discrepâncias	64

Exames complementares	67
Discrepâncias maiores	71
Concordância inter-observadores.....	73
Pacientes vulneráveis	74
Considerações Finais	77
 CONCLUSÕES.....	83
 REFERÊNCIAS BIBLIOGRÁFICAS	86
 ANEXO DE PUBLICAÇÃO	100

Introdução

A falibilidade do processo diagnóstico

O processo diagnóstico na medicina traduz-se por uma complexa interação entre conhecimentos, habilidades e procedimentos técnicos em condições de incerteza (KASSIRER, 1995). As variáveis que o cercam são inúmeras e de difícil caracterização. Estes aspectos dificultam o desenvolvimento de estudos mais aprofundados sobre o processo de elaboração do diagnóstico e suas variáveis correlatas, apesar da tese largamente aceita que a precisão no diagnóstico é fundamental para o adequado cuidado ao paciente e dos enormes danos que o paciente e mesmo o médico podem sofrer quando este não é corretamente realizado.

A discussão sobre erros no sistema de cuidado à saúde, apesar de incipiente, vem mudando de um tema considerado tabu para um assunto cada vez mais freqüente na literatura médica (DESKIN; HOYE, 2004; MCNUTT; ABRAMS; ARON, 2001; WEINGART et al., 2000). Vários artigos e editoriais de revistas científicas de grande alcance têm colocado a questão dos erros no trabalho médico de forma mais constante e sistematizada em nossa literatura especializada (BRENNAN et al., 1991; LEAPE, 1994; REASON, 2000; TEMPLE, 2004). Este debate inicial é muito importante e fundamental para que se crie uma cultura diferente de se lidar com este delicado assunto.

O erro no diagnóstico médico dentro do sistema de saúde é encarado de forma vexatória e tratado como algo impossível de acontecer, e os profissionais envolvidos muitas vezes são considerados como criminosos. Obviamente, sabemos que todos os sistemas são passíveis de

erros e, se houver uma conscientização sobre isto, daremos o primeiro passo no sentido da avaliação e minoração destes problemas (LESTER; TRITTER, 2001).

No entanto, o grande desenvolvimento tecnológico dos últimos anos trouxe para a população em geral uma alta expectativa em relação à performance do trabalho médico como um todo (SANDERS, 2004). Este fato aliado à impressão anterior, sacerdotal, de que o médico deveria ser imune a erros, faz com que a visão geral sobre estes possíveis eventos seja extremamente negativa.

A própria classe médica sempre apresentou uma tendência de forte recriminação sobre os erros descobertos e muitas vezes de acobertamento de outros não evidentes (O'NEIL et al., 1993). Estas características fizeram com que não houvesse um tratamento mais racional e produtivo sobre o tema. O modo de abordagem deste problema em outras áreas do conhecimento que lidam com vidas humanas, como na aviação e na engenharia, foi mais rápido e focado na melhoria dos próprios serviços (BURKE et al., 2004; HELMREICH, 2000; SPENCER, 2000).

Uma perspectiva sistêmica dos erros clínicos propõe que cada erro é uma cadeia de circunstâncias envolvendo múltiplos atores dentro de uma organização como um todo (LEAPE et al., 1998). Uma abordagem atual é deslocar uma culpa individual para um entendimento ou concepção mais institucional. A consequência positiva é que os médicos podem examinar seus próprios erros sem envergonharem-se demasiadamente, visto que este tipo de sentimento acarreta dificuldades pessoais e da equipe em corrigi-los e prevenir futuros acontecimentos. Esta perspectiva sistêmica não deve, no entanto, reduzir o necessário grau de vigilância que os médicos devem manter para seu próprio monitoramento (BORREL-CARRIÓ; EPSTEIN, 2004).

Desta forma, a literatura médica passa a preocupar-se de maneira mais incidente com o estudo e a avaliação de erros no sistema de saúde. A partir de uma avaliação mais sistematizada, procurou-se estabelecer os níveis de atenção à saúde em que haveria uma capacidade de localização de possíveis erros e prevenção. Nesta tentativa de sistematização, houve uma divisão de estudos focados em diversas etapas dos cuidados à saúde (BATES, 2000; WIEMAN; WIEMAN, 2004).

Assim começa uma avaliação de pontos do sistema de cuidado à saúde em que há possibilidade de erros que poderiam ser mais facilmente distinguíveis e, portanto, com maior probabilidade de intervenção (WOLF et al., 2004). Ao final da década de 90 houve um direcionamento de vários estudos para a avaliação de erros na prescrição de medicamentos (BOBB et al., 2004; RIDLEY; BOOTH; THOMPSON, 2004). Estas avaliações observavam desde o ato de prescrição da medicação pelo médico, a maneira como esta foi passada à equipe de enfermagem (de forma escrita ou verbal), a capacidade de entendimento desta equipe, o correto preparo, a possibilidade de avaliação pela equipe da farmácia, a correta administração, entre outras (BARBER, 2004; GORDON, 2001; KARSH, 2004). Com o levantamento destes dados, propostas de alterações nestes mecanismos foram feitas como prescrições eletrônicas, avaliação do farmacêutico, pessoalmente ou através de programas computadorizados como, por exemplo, compatibilidade de peso *versus* dose *versus* interações medicamentosas, programas de verificação de eventos adversos para o uso de medicamentos, entre outros (KUPERMAN; GIBSON, 2003).

Outra forma encontrada para avaliar o problema foi através de análise de sistemas, definindo o trabalho de cuidado à saúde como um sistema de múltiplas variáveis e criando

sistemas preventivos de erros, através de análise direta ou via programas computadorizados, alimentados com algumas variáveis passíveis de mensuração (DENISON; PIERCE, 2003).

Por ser um assunto discutido de maneira mais abrangente e de forma menos emocional na literatura médica internacional, só recentemente podemos observar diversos enfoques que puderam ser obtidos. A sistematização destas abordagens é benéfica e salutar para o próprio progresso do estudo sobre este assunto de forma mais organizada e dividida em partes. As medidas e os sistemas idealizados para uma possível diminuição de erros no cuidado à saúde, provenientes desta divisão, podem ser úteis e devem ser testadas e implementadas.

No entanto, observa-se que as abordagens sobre estes possíveis erros, expostas anteriormente, são insuficientes para a investigação sobre erros cognitivos no processo de elaboração do diagnóstico. Os estudos sobre o processo diagnóstico são pouco representativos em número e não estimularam a comunidade científica a aprofundá-los, mesmo sabendo que o correto diagnóstico é essencial para o adequado manejo e tratamento de qualquer agravo à saúde. Uma revisão da coluna intitulada “Uses of error” da revista Lancet, por exemplo, detectou que 52% dos casos de erros descritos estavam relacionados com o erro diagnóstico (DUNN; REDDY; BOWES, 2005).

Esta avaliação mais sistemática de possíveis erros no processo de tomada de decisões foi mais bem estudada na literatura sobre a psicologia (SUTHERLAND, 2002). Muitas decisões são baseadas na opinião e na crença das pessoas, subsidiadas por conhecimentos prévios que podem ser de vivência ou adquiridos em teoria. No momento da tomada de decisão os modos psicológicos que auxiliam na resolução de problemas (heurística) podem ser enviesados pelas incertezas do momento (TVERSKY; KAHNEMAN, 1974).

A partir de uma abordagem mais psicológica da elaboração de um diagnóstico, alguns autores passaram a tentar correlacionar erros com os passos do processo de elaboração do diagnóstico. Voytovich, Rippey e Suffredini (1985) começaram a estudar e classificar os possíveis modos de erro no raciocínio diagnóstico, classificando-os, a princípio, em quatro categorias. A primeira seria a omissão, onde um importante indício ou pista diagnóstica seria simplesmente ignorado. A segunda seria o encerramento prematuro do diagnóstico, onde o diagnóstico assumido para o paciente não corresponde a todos os dados apresentados por ele. A terceira seria uma síntese errada, quando os dados disponíveis contradizem a conclusão. A quarta seria uma síntese inadequada, onde as conclusões diagnósticas que seriam possíveis com os dados apresentados não são realizadas. Neste estudo os autores apresentaram três casos clínicos a diversos médicos e tentaram categorizar os erros de diagnóstico, quando cometidos. O encerramento prematuro do diagnóstico foi o erro mais freqüente.

Kassirer e Kopelman (1989) identificaram quatro estágios para o estabelecimento de um diagnóstico que eles adaptaram a partir do sistema de resolução de problemas da ciência cognitiva moderna, que segue o processo de descoberta científica. O primeiro seria o desencadeamento (*“triggering”*) de uma ou mais possibilidades diagnósticas, que trata do processo de estabelecer estas hipóteses a partir das informações iniciais de anamnese e exame físico obtidos. O segundo seria o desenvolvimento de uma representação cognitiva, que envolve a formulação (*“framing”*) do problema diagnóstico dentro de um contexto que determinará a direção das investigações futuras. O terceiro seria a reunião e o processamento (*“processing”*) das informações, que inclui a coleta de informações adicionais levando ao desenvolvimento e revisão das hipóteses trabalhadas. O quarto e último estágio seria a verificação (*“verification”*),

que envolve a nomeação do diagnóstico que foi estabelecido com um elevado grau de probabilidade.

A partir destes estágios os autores propuseram uma identificação e classificação para obter um aprofundamento no conhecimento da epidemiologia e causas dos erros cognitivos. Esta classificação também é derivada da literatura científica, assumindo que a base da resolução de problemas clínicos e o processo cognitivo da descoberta científica são intrinsecamente similares. A proposta caracteriza-se por cinco itens. O primeiro seria o desencadeamento defeituoso, o segundo a formulação de contexto defeituosa, o terceiro seria o defeito na coleta e no processamento de informações, o quarto seria a falta de verificação e o quinto seria a presença de erros sem defeitos.

Esta linha de estudos não obteve grande adesão, incitando a produção de outros poucos trabalhos científicos na tentativa de elaborar uma sistematização dos erros no processo diagnóstico e, principalmente, na utilização de outras metodologias para esta avaliação (MARGO, 2003).

A necropsia e as discrepâncias diagnósticas

Os achados de necropsia sempre foram usados como substrato para averiguação das doenças do paciente que morreu, sendo, portanto, uma peça de avaliação do diagnóstico clínico que vem sendo utilizada há muito tempo. Vários investigadores já demonstraram a grande validade da necropsia na auditoria dos diagnósticos, sendo aceito de forma generalizada que é

uma ferramenta importante e indispensável para a análise das discrepâncias dos diagnósticos clínicos (LARDENOYE et al., 2002; NADROUS et al., 2003; NEWTON et al., 2004). Apesar do fato de que a necropsia deve também ser objeto de controle de qualidade para assegurar a adequada interpretação clínica dos achados morfológicos, incluindo os diagnósticos realizados pelo patologista em seu relatório final de necropsia, esta se mantém como o padrão ouro para a avaliação do diagnóstico clínico (BELZEBERG; RIVKIND, 1998; SHOJANIA; BURTON, 2004; TRELSTEAD et al., 1996).

Estudos de comparação entre os diagnósticos que foram atribuídos clinicamente pelo médico assistente após a morte do paciente com os diagnósticos verificados pelo patologista após a necropsia iniciam-se em 1912. Cabot (1912) apresentou uma análise de 3000 exames necroscópicos realizada no *Massachusetts General Hospital*, com um percentual de 40% de discordâncias entre os diagnósticos clínicos e os achados de necropsia. Na primeira metade do século XX, a análise comparativa de alguns estudos demonstrou resultados semelhantes em relação às discordâncias diagnósticas (GALL, 1968).

Na década de 80, Goldman et al. (1983) realizaram um estudo utilizando uma nova metodologia de análise destas discrepâncias diagnósticas e, por meio dela, propuseram uma nova classificação para as mesmas. Neste estudo, eles compararam as discrepâncias entre os diagnósticos clínicos e necroscópicos ocorridas em três diferentes anos, separados cada um deles por uma década (1960, 1970 e 1980). Demonstraram não haver tendência de melhora na correlação diagnóstica no decorrer deste intervalo de vinte anos estudado. Esta nova classificação separou as discrepâncias em quatro categorias, sendo que as duas primeiras foram descritas como discrepâncias maiores e as últimas como discrepâncias menores.

A classe I refere-se aos diagnósticos relacionados com a causa básica ou direta do óbito que, caso fossem realizados anteriormente, poderiam ter mudado a sobrevida do doente. A classe II refere-se a estes diagnósticos que, mesmo se realizados anteriormente, não trariam a certeza de alguma melhora na sobrevida e/ou alguma modificação no tratamento do doente. Estas duas classes foram denominadas de discrepâncias maiores, por estarem justamente relacionadas com as causas diretas ou com a causa básica do óbito. As discrepâncias menores consistiam em diagnósticos secundários que não apresentavam relação com as causas do óbito, porém significavam doenças que deveriam ser tratadas (classe III) ou que estavam em estágio incipiente (classe IV) e, portanto, de difícil diagnóstico.

Diversos estudos nos últimos trinta anos, utilizando-se da necropsia como padrão ouro para a confirmação da concordância diagnóstica, não demonstraram um aumento significativo da mesma (BATTLE et al., 1987; BAYER-GARNER; FINK; LAMPS, 2002). Apresentamos no quadro 1 alguns estudos publicados nos últimos dez anos sobre as concordâncias diagnósticas com base na necropsia.

Estes estudos baseados na necropsia mostram que em várias populações, grupos de doenças, pacientes de diferentes faixas etárias ou em ambientes hospitalares específicos (ERMENC, 2000; GOLDSTEIN et al., 1996), as discrepâncias diagnósticas continuam a ocorrer em percentuais razoavelmente elevados, com algumas variações particularizadas (LANDEFELD et al., 1988; SALIB; TADROS; AMBROSE, 2000). Não é possível realizar comparações diretas entre estudos de circunstâncias tão diversas como as apresentadas no quadro 1, no entanto, é demonstrado que os trabalhos baseados no exame necroscópico apontam as discrepâncias diagnósticas de forma inequívoca, em diversos ambientes e grupos de população.

Quadro 1: Trabalhos de correlação diagnóstica publicados nos últimos 10 anos

Autor	Ano	Necropsias (%)	Amostra (n)	Discordâncias (%)	Local do estudo
SEOW; LAU	1996	49,06	236	T-23,2	Sala de emergências
DHAR et al.	1998	62	336	M-18 T-58	UTI neonatal
KUMAR et al.	1998	36	107	M-13 T-34	Hospital infantil
THORNTON; O'HARA	1998	47,4	174	T-28,2	Óbitos perinatais
ERMENC	1999a	-	921	T-36,2	Hospital geral
MORT; YESTON	1999	29	149	M-23 T-41	UTI cirúrgica
ZARBO; BAKER; HOWANITZ	1999	-	2479	M-39,7	Estud. multicêntrico
KUMAR et al.	2000	61	296	M-12 T-44	UTI neonatal
TWIGG; MCCRIRICK; SANDERSON	2001	40,5	97	T-23,7	UTI adulto
TAI et al.	2001	22,7	91	M-19,8	UTI adulto
ONG et al.	2002	97	153	M-19	UTI trauma
BAYER-GARNER; FINK; LAMPS	2002	20,14	142	T-49	Hosp. universitário
NADROUS et al.	2003	33	455	M-21	UTI adulto
*COMBES et al.	2004	53	167	M-31,7	UTI cirúrgica
GRADE et al.	2004	77	4491	T-32,1	Hosp. geral
NEWTON et al.	2004	40	61	M-20	Hospital pediátrico
SPILIOPOULOU et al.	2005	-	252	M-19	Pop. adulta
AALTEN; SAMSON; JANSEN	2006	28	93	M-39	Pac. geriátricos
BUCKNER; BLATT; SMITH	2006	-	533	M-20,5	Hospital pediátrico

M = discordâncias diagnósticas maiores; T = discordâncias totais

* estudo prospectivo

Anderson et al. (1989) analisaram as tendências de sensibilidade e de especificidade para o diagnóstico clínico de onze doenças, quando estavam relacionadas com o óbito das pessoas. Este estudo teve uma abrangência de cinco décadas (1930 - 1977) e mostrou uma tendência de melhora nos índices avaliados de sensibilidade diagnóstica, em apenas duas das doenças escolhidas no período. Egerváry et al. (2000), estudando neoplasias pulmonares, apresentaram um percentual elevado de casos falsos positivos (56%) e de falsos negativos (30%) para o diagnóstico clínico destes tumores após a avaliação diagnóstica pela necropsia, sem tendência de melhora com relação a estudos anteriores (SILVERBERG, 1984). Burton et al. (1998) acharam uma discrepância de 44% para os diagnósticos de neoplasias malignas, sem tendência de diminuição em relação a outros trabalhos. Zher et al. (1997) demonstraram que em um terço dos pacientes que faleceram após operações cardíacas, e que foram submetidos à necropsia, havia informações não detectadas clinicamente as quais poderiam ter alterado a terapia do doente. Portanto, pode-se observar que mesmo em estudos que consideram apenas doenças específicas relacionadas a óbitos, os índices de discrepâncias não vêm obtendo melhora.

Avaliando-se o quadro 1, nota-se que diversos estudos foram realizados em unidades de terapia intensiva. Foram estudados também grupos de pacientes pediátricos, neonatais e os óbitos perinatais. Em todos, as diferenças diagnósticas apresentam-se sempre em percentuais elevados, com variações particulares. Se considerarmos que estas discrepâncias podem corresponder em boa parte das vezes a erros diagnósticos que, se evitados levariam a mudanças nas formas de tratamento deste pacientes, pode-se inferir a importância destes estudos e suas implicações para a melhoria na qualidade do trabalho médico (LEAPE, 1994; O'LEARY, 1996). Outro fator importante é que mesmo os erros sem conseqüências para o prognóstico do doente acabam em incorrer em números equivocados nas estatísticas de saúde, principalmente

nas de mortalidade (TSUJIMURA et al., 1999; VALDEZ-MARTÍNEZ; ARROYO-LUNAGÓMEZ; LANDERO-LÓPEZ, 1998).

Roulson et al. (2005), em uma revisão e meta-análise de trabalhos que usaram a necropsia para avaliar as discordâncias diagnósticas, concluíram que de 1960 até a época de seu estudo houve pouca melhora na taxa geral de discrepâncias diagnósticas. Por outro lado, Shojania et al. (2003) apresentaram outra revisão sistemática da literatura analisando os dados de 45 artigos diferentes, englobando um período de quarenta anos. Eles analisaram estudos em diferentes locais, com populações diferentes, tipos de pacientes diversos e taxas de realização de necropsia diferentes. Com estes dados elaboraram uma projeção estatística que indicaria um resultado de redução de 19,4% por década para as discrepâncias maiores e 33,4% para as discrepâncias tipo I. A partir destes resultados eles elaboraram uma estimativa válida somente para os Estados Unidos da América, pois consideravam apenas trabalhos norte-americanos, do percentual de erros que seriam encontrados em hospitais de seu país correlacionado com alguns percentuais de realização de necropsias diferentes (tabela 1).

Considerando a estimativa como sendo correta, podemos calcular que ao analisar apenas 5% dos óbitos de um serviço, deixa-se de diagnosticar 86% de todos os erros maiores e 92% dos erros tipo I. Mesmo elevando a análise para 37% de necropsias, deixa-se 24% e 48% de perdas para as categorias acima mencionadas, respectivamente. Os próprios autores estimaram em cerca de 35000 óbitos nos EUA com erros tipo I, atualmente. Outro dado interessante desta revisão é que os estudos que foram realizados de forma prospectiva e os que apresentaram a associação entre patologistas e médicos clínicos realizando a classificação de erros apresentaram maior percentual de erros tipo I, não tendo, no entanto, discriminado com clareza estes números em sua revisão.

Tabela 1 – Estimativa de encontro de erros em hospitais americanos segundo o percentual de realização de necropsias

Taxa de realização necropsias	Discrepâncias Tipo I	Discrepâncias Maiores
100%	4,1%	8,4%
37%	5,8%	17,4%
5%	6,7%	24,4%

Os próprios autores discutem as possíveis limitações de sua revisão com diferentes locais, pacientes e taxas de necropsias. Embasam parte de suas considerações em um único estudo que manteve um percentual alto e fixo de realizações de necropsias, pois foi este que apresentou melhora na avaliação de três diferentes décadas (SONDEREGGER-ISELI et al., 2000).

Entretanto, mesmo com os resultados expressos acima os autores demonstram uma preocupação com o declínio no número de necropsias (SHOJANIA; BURTON, 2004), realizadas em boa parte dos hospitais, principalmente nos Estados Unidos da América e alguns países da Europa (BÖTTIGER, 1992; ERMENC, 1999b).

Esta tendência foi responsável pelo início de inúmeras discussões acerca do papel das necropsias na medicina moderna, principalmente na segunda metade do século XX (LUNDERBERG, 1996). As discussões na literatura procuraram identificar alguns dos fatores causais desta tendência de diminuição no número de exames necroscópicos (HABER, 1996). Portanto, houve um incremento nos estudos relacionados ao mesmo, associado a uma procura dos fatores relacionados com esta tendência (HUTCHINS, 1996a; SECKINGER, 1996).

Muitas razões têm sido propostas para explicar o declínio nas taxas de realização de necropsias. Diversos fatores de ordem religiosa (GELLER, 1984), social (KING; MEEHAN, 1973), política e econômica (LUNDBERG, 1984) foram abordados.

Mais especificamente, entre estes itens podemos citar: o não reembolso para uma tarefa que consome um grande tempo (CHERNOF, 1996); o medo de ações judiciais por erro médico (ANDERSON; HILL, 1988) mesmo com relatos de literatura confirmando o valor da necropsia na defesa contra processos judiciais (HAQUE; PATTERSON; GRAFE, 1996); o excesso de confiança no diagnóstico clínico (PELLEGRINO, 1996); os avanços nas tecnologias de diagnósticos por exames conforme descrito anteriormente (FEINSTEIN, 1996); a falta de conhecimento acerca do seu valor e dos seus benefícios; a falta de políticas públicas no sentido de manter um número mínimo de necropsias junto aos hospitais (SETLOW, 1996); o fato de muitas faculdades de medicina não exigirem dos estudantes que assistam um exame necroscópico para sua formação (HILL, 1996); o ensino médico centrado em estruturas submicroscópicas (HUTCHINS, 1996b); as preocupações acerca da transmissibilidade de doenças como a variante de Creutzfeldt-Jakob e a síndrome da imunodeficiência adquirida durante o procedimento; o envolvimento de pessoal não médico nas decisões acerca da continuidade das necropsias (BAYER-GARNER; FINK; LAMPS, 2002); a cultura religiosa das sociedades (BIERIG, 1996); a falta de engajamento de muitos patologistas por considerarem este exame um procedimento trabalhoso e sem retorno financeiro (KING, 1984; SINARD; BLOOD, 2001); a falta de comunicação entre a patologia e a clínica por desinteresse de ambos, para ter acesso aos resultados do exame e assim prover uma retro-alimentação positiva para o mesmo (ZARBO; BAKER; HOWANITZ, 1999); a delegação do exame necroscópico aos patologistas menos graduados (DIAMOND, 1996); a falta de treinamento da equipe de saúde no

modo de pedir permissão para a realização do exame e no contato com a família (MCPHEE, 1996a).

Dessa forma, pode-se observar que são inúmeras as considerações acerca das causas do declínio do número deste importante exame. No entanto, desde que Morgagni (1682-1771) começou a correlacionar, de maneira sistemática, as observações clínicas advindas do paciente com os achados anatômicos percebidos durante a necropsia, no final da primeira metade do século XVIII, ela passou a ter um papel fundamental no progresso do saber médico (HILL; ANDERSON, 1988; WISSLER, 1984). Várias doenças e seus mecanismos fisiopatológicos foram desvendados por este exame (KAJIWARA; ZUCOLOTO, 1991). A necropsia manteve sempre um papel de destaque no aperfeiçoamento da investigação e da pesquisa de entidades patológicas (KAUFMAN, 1996), e segundo o Conselho para Assuntos Científicos da Associação Médica Americana, apresenta grande potencial de contribuição para as necessidades modernas da pesquisa científica (AMERICAN MEDICAL ASSOCIATION - COUNCIL ON SCIENTIFIC AFFAIRS, 1987). Além disso, ajudou inúmeras vezes no avanço de medidas terapêuticas, através de suas observações pós-morte, assim como solucionando casos complicados de interesse judicial (CORDNER, 1993).

A literatura médica vem demonstrando a partir da segunda metade do século XX as potencialidades da necropsia (BARON, 2000), Hill e Anderson (1996), por exemplo, realizaram uma listagem parcial de doenças descobertas ou esclarecidas através da necropsia desde 1950. O quadro apresentado pelos autores demonstra mais de oitenta doenças. Entre elas citamos: complicações de cirurgias cardíacas; doença da membrana hialina; doença hepática induzida pela aflatoxina; elucidação de várias glomerulonefrites; enteropatia perdedora de proteína;

complicações do diabetes mellitus; adrenoleucodistrofia; síndrome do choque tóxico. Estes são apenas alguns exemplos, escolhidos aleatoriamente do trabalho, para exemplificação.

Além destas contribuições para o esclarecimento de diversas doenças, deve-se lembrar o importante papel desempenhado por este exame no auxílio às estatísticas de mortalidade (FERNANDEZ-SEGOVIANO; ESTEBAN, 1999; HASUO et al., 1989). Estes números são de extrema valia para o diagnóstico de saúde em uma comunidade, e com estes dados pode-se realizar um planejamento de ações para a melhoria das condições em saúde (MCPHEE, 1996b). Portanto, quanto mais próximo da realidade estiverem as estatísticas de mortalidade, ações em saúde mais adequadas para a melhoria da coletividade poderão ser promovidas (BLOSSER; ZIMMERMAN; STAUFFER, 1998). Vários artigos descrevem a não exatidão na avaliação das causas básicas de óbito nas Declarações de Óbitos, principalmente relacionadas às imprecisões ao se considerar apenas o diagnóstico clínico (CAMERON; MCGOOGAN, 1981; KIRCHER; NELSON; BURDO, 1985).

Em outra abordagem, observa-se de maneira cada vez mais constante na literatura a descrição da necessidade de se realizarem atividades e medidas para controle de qualidade do trabalho médico (GUT; FERREIRA; MONTENEGRO, 1999; JUVIN et al., 2000). Neste sentido, o exame necroscópico seria colocado em papel central, por suas características citadas anteriormente (CHUTE; BALLARD; NEMETZ, 1991; GAMBINO, 1984).

Contudo, mesmo que as projeções estatísticas que Shojania e colaboradores propuseram estejam corretas, isto é, com algumas melhoras nas discrepâncias diagnósticas ao longo das últimas décadas, estas se mostram muito discretas diante de tantos avanços tecnológicos obtidos no período. A tecnologia disponível para procedimentos diagnósticos evoluiu de maneira impressionante em diversas áreas do conhecimento médico.

No diagnóstico por imagens, evoluiu-se da radiografia comum para as imagens anatômicas da ressonância nuclear magnética. Temos como opções os exames radiológicos contrastados, as ultra-sonografias, a medicina nuclear e a tomografia computadorizada, entre outros. Na medicina laboratorial esta evolução percebe-se desde a celeridade com que obtemos resultados de exames gasométricos e de eletrólitos até as novas possibilidades de dosagens de inúmeros elementos e hormônios em várias espécies de materiais. Na imunologia, apresentou-se uma revolução dos testes diagnósticos em virtude dos ensaios imunoenzimáticos, além do aumento na cobertura de doenças imunopreveníveis. Na patologia, passou-se da microscopia óptica para a eletrônica, com a associação de outras técnicas como a imunofluorescência. Realizar uma listagem de todos os avanços tecnológicos alcançados nos últimos anos é muito difícil, e, portanto, deixamos as citadas acima a título de exemplificação.

Com isto apresenta-se uma questão para a prática médica que não vêm sendo apropriadamente discutida: Qual é o real impacto das modernas tecnologias diagnósticas para a melhoria da precisão diagnóstica, isto é, houve uma melhora substancial no processo diagnóstico com a incorporação destas novas tecnologias?

Obviamente que os avanços nas tecnologias diagnósticas nos permitem realizar um número maior e mais detalhado de diagnósticos com melhora na sobrevida e na qualidade de vida das pessoas. Mas a associação da melhora nas condições de vida, atribuída à boa alimentação, higiene, educação entre outras, com os avanços na saúde em virtude destes progressos nas ciências médicas, também são fatores que propiciaram estas melhorias verificadas, principalmente, em países desenvolvidos. Portanto a questão colocada refere-se ao processo da elaboração do diagnóstico.

Objetivos

Realizar comparação analítica da concordância ou discordância diagnóstica entre três fases deste processo: diagnóstico clínico inicial, quando da entrada do paciente no serviço; diagnóstico clínico final, realizado após o óbito do paciente; e diagnóstico necroscópico, expresso no relatório de necropsia, durante os anos de 1980, 1990 e 2000.

Avaliar a suficiência da análise macroscópica da necropsia na investigação de achados discordantes dos diagnósticos clínicos, correlacionando-os com os diagnósticos necroscópicos completos, incluindo a microscopia.

Avaliar a precisão (“*accuracy*”), a sensibilidade e a especificidade diagnósticas para os capítulos mais prevalentes da Classificação Internacional de Doenças 10^a revisão (ORGANIZAÇÃO MUNDIAL DE SAÚDE, 1989).

Confrontar a concordância ou discordância diagnóstica de um grupo de pacientes considerados vulneráveis contra os demais.

Interpretar a adequação dos exames complementares nos pacientes com discrepâncias maiores.

A partir dos dados obtidos, proceder a uma análise crítica do processo diagnóstico, considerando os seus aspectos de precisão e falibilidade, assim como as considerações bioéticas disto decorrentes.

Material

Local do estudo

O Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo (HCFMRPUSP) é um hospital universitário geral e predominantemente público. Durante o período de estudo, realizava atendimentos à população nos três níveis de complexidade, primário, secundário e terciário. Apresenta uma unidade localizada na região central do município de Ribeirão Preto, destinada principalmente ao atendimento de urgências e emergências e por isso denominada Unidade de Emergência. Outra unidade, localizada no Campus da Universidade de São Paulo, destinada ao atendimento de casos de diversas especialidades médicas, e onde estão localizados os ambulatórios e enfermarias destas especialidades. Esta unidade apresenta, geralmente, casos de maior complexidade.

O HCFMRPUSP é referência para a região de Ribeirão Preto e todo o nordeste do estado. Portanto, além de atender à população de Ribeirão Preto, atende também pacientes vindos de sua regional de saúde, que é uma subdivisão geográfica do estado de São Paulo para a administração da saúde, Diretoria Regional de Saúde (DRS XVIIIIG). No entanto, ele recebe inúmeros pacientes provenientes de outras regionais de saúde do estado de São Paulo e de outros estados da federação, inclusive pacientes de outras regiões que não a sudeste.

O município de Ribeirão Preto contava com uma população de 316.918 habitantes em 1980, 434.142 em 1991 e 504.923 em 2000, segundo a Fundação SEADE e o Censo 2000 do Instituto Brasileiro de Geografia e Estatística (IBGE). Tem como principais atividades

econômicas a agroindústria de açúcar e álcool combustível e o setor de serviços. Apresenta setor universitário bem desenvolvido com três universidades e outros centros de ensino superior.

A DRS XVIII G, correspondente à região de Ribeirão Preto, é composta por 25 municípios, contabilizando 1.109.034 habitantes em sua área de cobertura, segundo o censo 2000 do IBGE.

A população total coberta pelo HCFMRPUSP é de difícil caracterização, visto que uma porcentagem importante de pacientes, atendidos pelo mesmo, vem de outros estados da federação, conforme já referido, principalmente do estado de Minas Gerais, mas até mesmo de outras regiões.

Pacientes

O universo abrangido por este estudo foi composto por todos os pacientes falecidos no HCFMRPUSP nos anos de 1980, 1990 e 2000, que foram submetidos ao exame necroscópico no Serviço de Patologia do Hospital, excluindo-se os óbitos fetais.

Fontes de dados

Os relatórios anuais de atividades do HCFMRPUSP de 1980 a 2000 forneceram diversas informações como: número de internações, taxa de mortalidade hospitalar e porcentagem de

pacientes encaminhados para necropsia, entre outras. Eles encontram-se disponíveis na assessoria técnica do HCFMRPUSP.

O HCFMRPUSP possui Serviço de Arquivo Médico (SAME) que mantém os prontuários completos dos pacientes internados em seus serviços. Os prontuários dos pacientes que foram amostrados para o estudo foram recuperados dos arquivos para análise e revisão de cada caso.

Os relatórios de necropsia contêm a identificação do cadáver, a causa de óbito fornecida após o exame macroscópico, colocada na Declaração de Óbito. Além disto, apresentam os relatos dos exames complementares e microscópicos quando realizados. Por último, trazem a causa do óbito após a correlação destes dados com a análise macroscópica.

Comitê de Ética

O estudo foi aprovado pelo Comitê de Ética em Pesquisa do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo em 20 de agosto de 2001, sob o número de protocolo 5724/2001.

Métodos

Amostragem

A amostragem foi realizada de maneira aleatória, dentro do universo de pacientes abrangidos pelo estudo, selecionando-se 300 indivíduos distribuídos igualitariamente para os anos de 1980, 1990 e 2000, de acordo com estudos anteriores da literatura (GOLDMAN et al., 1983; SONDEREGGER-ISELI et al., 2000).

Coleta dos dados

A coleta dos dados baseou-se na revisão e análise de cada prontuário médico amostrado para o estudo. Foram coletados os seguintes dados:

- Pessoais - idade, gênero, etnia;
- Hospitalares - número de admissões hospitalares nos últimos 12 meses antes da morte, tempo de estadia durante a internação onde ocorreu o óbito, diagnósticos clínicos e necroscópicos;
- Exames subsidiários em número, discriminação e diagnóstico do laudo - procedimentos radiológicos não contrastados, procedimentos radiológicos contrastados, endoscopias, biópsias, explorações cirúrgicas, cintilografias, ultra-sonografias, tomografias computadorizadas e ressonâncias nucleares magnéticas.

Os diagnósticos identificados durante esta análise foram inicialmente agrupados em quatro grupos, nomeados da seguinte forma: diagnósticos necroscópicos finais, diagnósticos necroscópicos iniciais, diagnósticos clínicos finais e diagnósticos clínicos iniciais.

Os diagnósticos necroscópicos finais foram retirados do relatório final de necropsia elaborado pelo Serviço de Patologia do HCFMRPUSP. Estes compreendem a análise do prontuário médico associada aos achados patológicos macroscópicos durante o procedimento e, quando necessário, a realização de exame histopatológico (microscopia) e de outros exames auxiliares à mesma. Por apresentarem esta abrangência de análises foram considerados como o padrão ouro de avaliação deste estudo.

Os diagnósticos necroscópicos iniciais foram retirados do relatório preliminar de necropsia. Este apresenta os diagnósticos listados inicialmente pelos patologistas e que foram encadeados na Declaração de Óbito do paciente.

Os diagnósticos clínicos finais foram coletados do prontuário médico do paciente, através de sua análise completa. Foram considerados os diagnósticos transcritos na evolução médica, na lista de problemas, na epicrise e na requisição de necropsia. Além disso, foram considerados diagnósticos aqueles assumidos como tal por terem levado a uma terapia específica, como no caso de um paciente em terapia intensiva que apresentou radiografia de tórax descrita na evolução com velamento e que foi iniciada terapia com antibióticos, mesmo não tendo sido transcrita a palavra pneumonia no prontuário.

Os diagnósticos clínicos iniciais também foram coletados do prontuário médico do paciente. Foram considerados os diagnósticos transcritos na entrada do paciente, no caso novo, na folha de triagem, na ficha de atendimento na unidade de emergência. Foram os diagnósticos realizados sem o apoio de exames de imagem dentro do hospital, ou aqueles que foram

suspeitados durante a internação levando a uma investigação armada. Portanto, consideramos para tal também os diagnósticos diferenciais não descritos no prontuário. Como exemplo, temos uma úlcera péptica confirmada à endoscopia digestiva alta para uma hipótese inicial escrita de gastrite apenas.

Crítérios de exclusão

Este estudo não apresentou qualquer critério de exclusão dos prontuários amostrados de maneira aleatória.

Categorias comparativas

Os diagnósticos coletados foram agrupados em três fases distintas para comparação. Diagnóstico clínico inicial, diagnóstico clínico final e diagnóstico necroscópico. Este último grupo foi formado pelo diagnóstico necroscópico inicial e o final, visto que nem todos os pacientes apresentavam as duas categorias separadamente à análise dos prontuários.

Como houve o agrupamento em três fases diferentes de diagnóstico para cada paciente, foi necessário agrupar, para comparação de concordância, duas fases por vez. Desta forma, pôde-se constituir três categorias para esta análise.

Na categoria 1, foram pareados o diagnóstico clínico final e o diagnóstico necroscópico. Na categoria 2, foram pareados o diagnóstico clínico inicial e o diagnóstico necroscópico. Por fim, na categoria 3 do estudo, foram pareados o diagnóstico clínico inicial e o diagnóstico clínico final.

Classificação

A comparação entre as categorias descritas foi realizada de acordo com a classificação de discrepâncias diagnósticas descrita por Goldman et al. (1983) modificada por Battle et al. (1987) e por Sonderegger-Iseli et al. (2000). Ela é dividida em seis classes da seguinte forma:

Discrepâncias maiores

Classe I

Discrepâncias em diagnósticos maiores. O conhecimento antes do óbito teria levado a mudanças no manejo do doente que poderiam prolongar ou curar o paciente (ex.: infarto pulmonar tratado como pneumonia ou infecção fúngica tratada como bacteriana).

Classe II

Discrepâncias em diagnósticos maiores que se tivessem sido detectadas antes do óbito não teriam levado a aumento de sobrevida, mesmo com o tratamento adequado (ex.: insuficiência cardíaca biventricular secundária a estenose aórtica com uma embolia pulmonar não

diagnosticada, sepsis bacteriana corretamente tratada devido a uma osteomielite cervical pós-operatória não diagnosticada, em um paciente com artrite reumatóide.)

Discrepâncias menores

Classe III

Discrepâncias em diagnósticos menores não diretamente relacionados com o óbito, mas com sintomas que deveriam ser tratados ou que poderiam afetar o prognóstico do paciente (ex.: carcinoma pulmonar em um paciente com aneurisma de aorta infra-renal roto).

Classe IV

Discrepâncias em diagnósticos menores (ocultos), mas com possível importância epidemiológica ou genética (ex.: litíase biliar ou carcinoma *in situ*).

Concordância

Classe V

Diagnósticos não discrepantes.

Casos não classificados

Classe VI

Pacientes que morreram imediatamente após a admissão, sem procedimentos diagnósticos. Necropsia foi insatisfatória sem dados esclarecedores e diagnósticos não foram estabelecidos.

A classificação das discrepâncias, entre as categorias descritas, foi realizada por dois pesquisadores (um clínico e um patologista), após discussão de cada caso, em apenas um tipo de

classe. Os casos considerados como discrepâncias maiores foram revisados novamente para nova classificação, se necessário. Um terceiro pesquisador fez uma revisão de uma amostra de 10% dos casos selecionados para avaliação da concordância inter-observadores, do método de classificação utilizado.

Todos os casos foram classificados pela causa básica de óbito, tendo como base o diagnóstico final de necropsia, de acordo com a Classificação Internacional de Doenças – 10ª Edição (ORGANIZAÇÃO MUNDIAL DE SAÚDE, 1989).

Com o intuito de realizar uma comparação de precisão (“*accuracy*”), sensibilidade e especificidade diagnósticas para cada capítulo da CID 10, realizamos um agrupamento dos diagnósticos junto a cada um de seus 21 capítulos. Para cada categoria do estudo descrita anteriormente realizou-se esta divisão. O quadro 2 mostra a nomenclatura dos 21 capítulos da CID 10.

Como a classificação nos diversos capítulos da CID 10 deu-se pela causa básica de óbito, e nem sempre está foi a responsável pelas discrepâncias maiores encontradas (as causas diretas podem ser responsáveis por discrepâncias maiores), foi construída uma tabela descritiva dos diagnósticos que levaram às discrepâncias maiores.

Formando uma nova variável para análise, separamos os pacientes em um grupo que nomeamos de vulneráveis. Os pacientes vulneráveis foram definidos como aqueles que apresentavam uma das condições descritas a seguir: idade maior que 75 anos, etilistas crônicos, com seqüelas neurológicas ou com malformações congênitas do sistema nervoso. Os pacientes não-vulneráveis foram aqueles que não se enquadraram nesta definição.

Quadro 2: Capítulos do Manual de Classificação Estatística Internacional de Doenças, Lesões e Causas de Óbito (CID), 10ª revisão (1989)

Capítulo da CID	Nomenclatura
1	Algumas doenças infecciosas e parasitárias
2	Neoplasias
3	Doenças do sangue e dos órgãos hematopoéticos e alguns transtornos imunitários
4	Doenças endócrinas, nutricionais e metabólicas
5	Transtornos mentais e comportamentais
6	Doenças do sistema nervoso
7	Doenças dos olhos e anexos
8	Doenças do ouvido e da apófise mastóide
9	Doenças do aparelho circulatório
10	Doenças do aparelho respiratório
11	Doenças do aparelho digestivo
12	Doenças da pele e do subcutâneo
13	Doenças do sistema osteomuscular e do tecido conjuntivo
14	Doenças do aparelho geniturinário
15	Gravidez, parto e puerpério
16	Algumas afecções originadas no período perinatal
17	Malformações congênitas, deformidades e anomalias cromossômicas
18	Sintomas, sinais e achados anormais de exames clínicos e de laboratório não classificados em outra parte
19	Lesões, envenenamentos e algumas consequências de causas externas
20	Causas externas de morbidade e mortalidade
21	Fatores que influenciam o estado de saúde e o contato com os serviços de saúde

Por fim, ao estudar os pacientes que apresentavam discrepância classe I para a categoria de comparação 1 (diagnóstico clínico final $\times$ necroscópico), resolvemos avaliar os exames complementares realizados. De acordo com o diagnóstico necroscópico final, procuramos observar se os exames complementares que poderiam sugerir o diagnóstico estavam adequadamente requisitados, se quando os exames adequados foram realizados, sua interpretação pode ter sido errônea, se os exames adequados apresentaram resultados falso-negativos por baixa sensibilidade ou falso-positivos por falta de especificidade, ou se não era possível classificar estes casos nas três categorias descritas acima. Portanto realizamos uma nova categorização para comparação dos dados deste grupo (classe I) especificamente.

A figura 1 apresenta o desenho do estudo.

Análise estatística

Os dados gerais dos pacientes foram expressos através de média, mediana e percentuais. Estas variáveis foram analisadas pelo teste não-paramétrico de Mann-Whitney comparando-se o grupo de discrepâncias maiores segundo Goldman et al. com o restante dos pacientes, à exceção da variável gênero que foi avaliada pelo teste do qui-quadrado.

Os resultados de discrepâncias diagnósticas foram analisados com o teste exato de Cochran-Armitage para tendências durante os anos de estudo, com $p \leq 0,05$.

A avaliação da concordância inter-observadores pela amostragem (10%) avaliada foi realizada com o cálculo do coeficiente de concordância Kappa, com a seguinte interpretação do

poder de concordância: $< 0,20$ = pobre; $0,21 - 0,40$ = regular; $0,41 - 0,60$ = moderado; $0,61 - 0,80$ = bom; $0,81 - 1,00$ = muito bom (ALTMAN, 1991).

Na análise do nível de concordância, entre observadores diferentes ou entre métodos de abordagem diversos, o coeficiente de concordância κ (kappa) é um indicador mais adequado (COHEN, 1960). Sua metodologia baseia-se na comparação entre as proporções de casos de concordância observados com a proporção de concordância que seria esperada ao acaso.

Foram calculadas a precisão (“*accuracy*”), a sensibilidade e a especificidade para os capítulos da CID com o maior número de casos: Doenças do aparelho circulatório, Neoplasias, Doenças do aparelho digestivo, Doenças do aparelho respiratório, Algumas doenças infecciosas e parasitárias, e Transtornos mentais e comportamentais. Estes foram realizados de acordo com as seguintes definições e fórmulas:

Verdadeiro positivo é um caso que se manteve sem discrepância no capítulo em avaliação, após o exame.

Verdadeiro negativo é um caso que se manteve discrepante do capítulo em avaliação, após o exame.

Falso positivo é um caso que havia discrepância maior com o diagnóstico clínico estando no capítulo da CID em estudo e o diagnóstico necroscópico em outro ou quando havia discrepância maior estando as duas fases no mesmo capítulo da CID.

Falso negativo é um caso em que havia discrepância maior com o diagnóstico necroscópico estando no capítulo da CID em estudo e o diagnóstico clínico em outro ou quando havia discrepância maior estando as duas fases no mesmo capítulo da CID.

Precisão foi calculada como soma de verdadeiros positivos e verdadeiros negativos divididos por todos os casos. [Precisão (%) = $VP+VN / \text{Todos}$]

Sensibilidade foi a proporção de verdadeiros positivos divididos por verdadeiros positivos mais os falsos negativos. [Sensibilidade (%) = $VP / (VP + FN) * 100$]

Especificidade foi a proporção de verdadeiros negativos divididos pela soma de verdadeiros negativos e falsos positivos. [Especificidade (%) = $VN / (VN + FP) * 100$]

Onde: VP - verdadeiro positivo

FP - falso positivo

VN - verdadeiro negativo

FN - falso negativo

Figura 1 – Desenho do estudo

Resultados

Em 1980 o HCFMRPUSP contava com 519 leitos em atividade, realizou 18225 internações durante o ano, onde ocorreram 597 mortes, sendo que 543 foram submetidos à necropsia no Serviço de Patologia do hospital, perfazendo 90,9% destes pacientes. A população do estudo, portanto, foi realizada com uma amostra de 16,7% dos óbitos e 18,4% dos submetidos à necropsia. A amostra foi composta por 45% de pacientes do gênero feminino e 55% do masculino. A idade média dos pacientes foi de 47,39 anos e a mediana foi de 51,5, com a idade mínima de 2 meses e a máxima de 99 anos. Em relação à cor da pele descrita no prontuário, tivemos 71% relatados como cor branca, 16% de cor parda, 12% de cor preta e 1% de cor amarela. Os pacientes não tiveram admissões anteriores no hospital em 64% dos casos, tiveram uma admissão anterior em 24%, e duas ou mais admissões anteriores em 12%. O tempo médio de duração da estadia no hospital na internação do óbito foi de 16,14 dias, com mediana de 8,5 dias, e estadia mínima de um dia e máxima de 104 dias.

Em 1990 o HCFMRPUSP contava com 624 leitos em atividade, com 29359 internações durante o ano, onde ocorreram 1325 mortes, sendo que 989 foram submetidos à necropsia no Serviço de Patologia do hospital, perfazendo 74,6% destes pacientes. A população do estudo, portanto, foi realizada com uma amostra de 7,5% dos óbitos e 10,1% dos submetidos à necropsia. A amostra foi composta com 47% de pacientes do gênero feminino e 53% do masculino. A idade média dos pacientes foi de 47,16 anos e a mediana foi de 50,5, com a mínima de 9 dias e a máxima de 86 anos. Em relação à cor da pele descrita no prontuário, tivemos 76% relatados como cor branca, 14% de cor parda, 10% de cor preta e nenhuma de cor amarela. Os pacientes não tiveram admissões anteriores no hospital em 71% dos casos, tiveram uma admissão anterior em 12%, e duas ou mais admissões anteriores em 17%. O tempo médio

de duração da estadia no hospital na internação do óbito foi de 15,33 dias, com mediana de 6 dias, e estadia mínima de 1 dia e máxima de 164 dias.

Em 2000 o HCFMRPUSP contava com 817 leitos em atividade, com 34075 internações durante o ano, onde ocorreram 1479 mortes, sendo que 944 foram submetidos à necropsia no Serviço de Patologia do hospital, perfazendo 63,8% destes pacientes. A população do estudo, portanto, foi realizada com uma amostra de 6,8% dos óbitos e 10,6% dos submetidos à necropsia. A amostra foi composta com 53% de pacientes do gênero feminino e 47% do masculino. A idade média dos pacientes foi de 56,7 anos e a mediana foi de 65, com a mínima de 16 dias e a máxima de 97 anos. Em relação à cor da pele descrita no prontuário, tivemos 83% relatados como cor branca, 5% de cor parda, 12% de cor preta e nenhuma de cor amarela. Os pacientes não tiveram admissões anteriores no hospital em 60% dos casos, tiveram uma admissão anterior em 15%, e duas ou mais admissões anteriores em 25%. O tempo médio de duração da estadia no hospital na internação do óbito foi de 13,51 dias, com mediana de 7 dias, e estadia mínima de 1 dia e máxima de 194 dias.

As tabelas 2 e 3 mostram o conjunto dos dados gerais dos pacientes (gênero, idade, dias de internação, admissões anteriores) em cada ano e no total. O gráfico 1 apresenta a evolução de admissões anteriores nos anos do estudo.

Tabela 2 – Gênero, idade e dias de internação dos pacientes estudados.

Perfil dos pacientes	1980(n=100)	1990(n=100)	2000(n=100)	Total (n=300)
Masculino/Feminino (%)	55 / 45	53 / 47	47 / 53	51,6 / 48,4
Idade (mediana/média anos)	51,5 / 47,4	50,5 / 47,1	65 / 56,7	55 / 50,4
Internação (mediana/média dias)	8,5 / 16,1	6 / 15,3	7 / 13,5	7 / 15

Tabela 3 – Admissões anteriores no HCFMRPUSP.

Pacientes	1980(n=100)	1990(n=100)	2000(n=100)	Total (n=300)
Sem admissões anteriores (%)	64	71	60	65
1 admissão anterior (%)	24	12	15	17
≥ 2 admissões anteriores (%)	12	17	25	18

Gráfico 1 – Evolução de admissões anteriores no HCFMRPUSP, pelos anos do estudo.

Em nossa amostragem, o exame necroscópico completo que inclui a histopatologia (microscopia) foi realizado em 100% dos casos em 1980, 82% em 1990 e 55% em 2000. O exame histopatológico elucidou o caso, completou ou modificou os diagnósticos macroscópicos maiores em 7% dos casos de 1980, 4,9% de 1990 e 7,3% de 2000. Evidenciou ainda diagnósticos menores não suspeitados clinicamente em 20% dos casos em 1980, 21,9% em 1990 e 23,6% em 2000.

Ao classificarmos cada caso pela causa básica de óbito nos diversos capítulos da CID 10^a Revisão, encontramos as doenças cardiovasculares como o maior grupo de nossa amostragem, sendo responsáveis, no total, pelo maior número de óbitos (Tabela 4). As neoplasias formaram o segundo maior grupo, seguido das doenças do aparelho digestivo, respiratório, doenças infecciosas e transtornos mentais.

A tabela 5 apresenta a avaliação de precisão, sensibilidade e especificidade dos casos distribuídos pelos principais capítulos diagnósticos da CID 10^a Revisão de acordo com sua causa básica de óbito. As doenças cardiovasculares apresentaram discreta tendência de melhora na sua especificidade e precisão. As neoplasias não apresentaram alterações significativas. As doenças do aparelho digestivo tiveram uma diminuição discreta na sua precisão, mantendo uma sensibilidade baixa, em queda de 67% para 50%. As doenças respiratórias apresentaram queda em sua precisão de 100% para 95% e em sua sensibilidade de 100% para 70%. Os transtornos mentais apresentaram queda em sua precisão, sensibilidade e especificidade.

Tabela 4 – Distribuição dos diagnósticos de acordo com grupos clínicos.

Capítulos da CID 10 ^a Revisão	1980	1990	2000	Total %
Doenças do aparelho circulatório	43	33	30	35,3
Neoplasias	12	24	23	19,7
Doenças do aparelho digestivo	6	9	6	7
Doenças do aparelho respiratório	3	5	10	6
Algumas doenças infecciosas e parasitárias	10	2	5	5,7
Transtornos mentais e comportamentais	5	5	7	5,7
Malformações congênitas, deformidades e anomalias cromossômicas	4	3	6	4,3
Doenças endócrinas, nutricionais e metabólicas	3	6	2	3,7
Doenças do aparelho geniturinário	6	1	3	3,3
Doenças do sistema nervoso	2	4	4	3,3
Outros	6	8	4	6
Total	100	100	100	100

Tabela 5 – Precisão, sensibilidade e especificidade para grupos diagnósticos.

Grupo diagnóstico	Precisão			Sensibilidade			Especificidade		
	1980	1990	2000	1980	1990	2000	1980	1990	2000
Dças cardiovasculares	90%	91%	93%	78,5%	79%	76%	86%	90%	93%
Neoplasias	96%	96%	96%	75%	83%	83%	95%	95%	95%
Dças do ap. digestivo	98%	97%	96%	67%	67%	50%	98%	99%	98%
Dças respiratórias	99%	95%	95%	100%	100%	70%	99%	95%	96%
Dças infecciosas	97%	99%	97%	70%	100%	80%	96%	99%	98%
Transtornos mentais	100%	96%	97%	100%	60%	67%	100%	99%	97%

O gráfico 2 apresenta a comparação entre o diagnóstico clínico final e o necroscópico nos três anos estudados. Pode-se perceber que não houve alteração nas classes I e II, discrepâncias maiores, no período. Houve uma diminuição das classes diagnósticas III e IV, discrepâncias menores. Com isto apresentou-se um nítido aumento da classe V, concordância diagnóstica, que era de 41 casos em 1980, 48 casos em 1990 e foi para 55 casos em 2000, tendo significância estatística ($p = 0,05$) ao longo das décadas.

Gráfico 2 – Distribuição da classificação de discrepâncias na comparação diagnóstico clínico final x necroscópico (*tendência de crescimento estatístico significativo $p = 0,05$)

O gráfico 3 apresenta a comparação entre o diagnóstico clínico inicial e o necroscópico. As discrepâncias maiores não apresentam tendência de melhora no período, apresentando-se com um valor mais elevado que na comparação anterior. Nota-se também um aumento na classe V, concordância diagnóstica, com significância estatística ($p = 0,03$), com a mesma diminuição concomitante nas discrepâncias menores, classes III e IV.

Gráfico 3 – Distribuição da classificação de discrepâncias na comparação diagnóstico clínico inicial x necroscópico (*tendência de crescimento estatístico significativo $p = 0,03$)

O gráfico 4 apresenta a comparação entre o diagnóstico clínico inicial e o final. Mostra-se evidente o alto percentual de concordância, classe V, nos três anos estudados. As discrepâncias maiores têm discreto aumento, enquanto as demais apresentam valores mínimos.

Gráfico 4 – Distribuição da classificação de discrepâncias na comparação diagnóstico clínico inicial x clínico final

Os exames ou procedimentos diagnósticos estão discriminados na tabela 6 de acordo com os anos de estudo e totais. As endoscopias, biópsias, ultra-sonografias, tomografias

computadorizadas e ressonâncias magnéticas tiveram aumento importante no período avaliado. Houve diminuição dos exames radiológicos contrastados e dos procedimentos cirúrgicos. Os exames radiológicos não contrastados e cintilografias mantiveram-se estáveis em número. O número total de exames e procedimentos subiu de 217 para 366 no período.

Tabela 6 – Número de exames ou procedimentos para os anos do estudo e percentual de concordância do laudo com achados de necropsia

Procedimentos	1980	1990	2000	Total
Radiológicos não contrastados	139 (96%)	118 (94%)	117 (96%)	374 (95%)
Radiológicos contrastados	39 (90%)	25 (100%)	14 (100%)	78 (95%)
Endoscopia	13 (92%)	10 (100%)	32 (97%)	55 (96%)
Biópsias	23 (96%)	43 (88%)	57 (98%)	123 (94%)
Exploração cirúrgica	8 (100%)	6 (100%)	3 (100%)	17 (100%)
Cintilografia	11 (100%)	9 (67%)	10 (100%)	30 (90%)
Ultra-sonografia	2 (50%)	35 (94%)	65 (98%)	102 (96%)
Tomografia computadorizada	1 (100%)	14 (100%)	59 (100%)	74 (100%)
Ressonância magnética	0 (--)	0 (--)	9 (100%)	9 (100%)
Total	236 (95%)	260 (93%)	366 (98%)	862 (96%)
Pacientes não submetidos a exames	6	6	2	14

Os diagnósticos que levaram à classificação de discrepâncias maiores, classe II e I, nos três anos estudados totalizaram 26 e 37 casos respectivamente. Apresentamos estes diagnósticos

em duas tabelas descritivas, 7 e 8. Nota-se que as doenças gastrintestinais formaram o maior grupo no tipo I, com destaque para 7 casos de úlcera péptica perforada. As pneumonias aparecem nas duas tabelas com 4 e 6 casos respectivamente.

Tabela 7 – Descrição e distribuição de diagnósticos das discrepâncias classe II discriminados por ano do estudo

Ano	Sistema	Nº	Doença e localização (n)
1980	Cardiovascular	4	Acidente vascular cerebral (1), valvulopatia mitral e aórtica (1), doença de Chagas (1), fibroelastose hipertrófica (1).
	Respiratório	3	Broncopneumonia (2), embolia pulmonar (1).
	Gastro-intestinal	3	Iatrogenia cirúrgica (1), peritonite (1), necrose intestinal maciça (1).
1990	Respiratório	3	Broncopneumonia (3).
	Cardiovascular	3	Acidente vascular cerebral (1), doença isquêmica cardíaca (1), cardiomiopatia dilatada (1).
	Gastro-intestinal	2	Desnutrição protéico-calórica (1), peritonite (1).
2000	Respiratório	5	Carcinoma broncogênico (1), fibrose pulmonar (1), doença pulmonar obstrutiva crônica (1), hemorragia pulmonar (1), embolia pulmonar (1).
	Cardiovascular	3	Acidente vascular cerebral (2), doença isquêmica cardíaca (1).

Tabela 8 – Descrição e distribuição de diagnósticos das discrepâncias classe I discriminados por ano do estudo

Ano	Sistema	n ^o	Doença e localização (n)
1980	Gastro-intestinal	5	Peritonite (3), úlcera duodenal perfurada (1), iatrogenia cirúrgica (1).
	Respiratório	3	Broncopneumonia (1), embolia pulmonar (1), pneumotórax (1).
	Cardiovascular	2	Infarto agudo do miocárdio (1), insuficiência cardíaca + edema agudo pulmão (1).
	Renal	1	Insuficiência renal aguda (1).
1990	Cardiovascular	7	Insuficiência cardíaca + edema agudo pulmão (3), endocardite bacteriana (1), infarto agudo do miocárdio (1), doença isquêmica (1), aneurisma dissecante de aorta (1).
	Gastro-intestinal	3	Úlcera péptica perfurada (3).
	Nervoso central	2	Meningite bacteriana (1), meningite tuberculosa (1).
	Respiratório	1	Corpo estranho (1).
2000	Gastro-intestinal	6	Úlcera péptica perfurada (3), Paracoccidioidomicose esofágica (1), iatrogenia cirúrgica (1), pancreatite (1).
	Respiratório	4	Broncopneumonia (2), hemorragia pulmonar (1), tuberculose pulmonar (1).
	Cardiovascular	3	Estenose aórtica (1), cor pulmonale (1), tamponamento cardíaco (1).

Os casos com discrepância classe I foram analisados em relação à adequação dos exames complementares que foram realizados. O gráfico 5 apresenta estes casos para os três anos do estudo e dos casos totais nas seguintes categorias de interpretações de adequação dos exames: 1 – exames que poderiam sugerir o diagnóstico não foram requisitados; 2 – exames apropriados foram realizados com uma possível interpretação equivocada; 3 – exames apropriados foram realizados, mas seus resultados foram falso-negativos por baixa sensibilidade ou falso-positivos por falta de especificidade; 4 – casos que não foi possível classificá-los nas categorias acima.

Gráfico 5 – Evolução numérica das categorias de interpretação da adequação dos exames, em pacientes com discrepâncias diagnósticas Classe I, e o total em percentual para cada categoria.

Analizamos estatisticamente as variáveis de idade, admissões anteriores, estadia, vulnerabilidade, exames comuns (definidos como os radiológicos, as endoscopias, as biópsias, as cintilografias e as explorações cirúrgicas), exames de imagem recentes (definidos como as ultra-sonografias, as tomografias computadorizadas e as ressonâncias magnéticas) e exames totais, na comparação de dois grupos: discrepâncias maiores x demais pacientes. Para tal utilizamos o teste não-paramétrico de Mann-Whitney. A variável gênero foi avaliada pelo teste do qui-quadrado. Estas comparações foram realizadas para os anos de 1980, 1990 e 2000 separadamente e no universo total dos casos estudados.

No ano de 1980 todas as variáveis descritas acima não apresentaram diferenças significativas na comparação dos dois grupos descritos. Para o ano de 1990 não ocorreram diferenças significativas para as variáveis idade, admissão, vulnerabilidade, exames comuns, de imagem recentes e totais. A variável gênero apresentou $p = 0,06$, sem diferença significativa, porém muito próximo do nível de significância adotado, com tendência das discrepâncias maiores serem significantes ao percentual de homens em relação às mulheres. A variável estadia apresentou $p = 0,05$, com diferença significativa, tendo o grupo dos demais pacientes possuindo valores superiores de estadia (média de $17,18 \times 8,38$ dias).

Na análise do ano 2000 as variáveis idade, gênero, estadia, vulnerabilidade, exames comuns, de imagem recentes e totais não apresentaram diferenças significativas na comparação dos dois grupos. A variável admissão teve $p = 0,007$, com o grupo de discrepâncias maiores tendo valores de admissões significativamente menores que o grupo demais pacientes (média de $1,29 \times 2,15$ admissões).

Em relação ao grupo total não houve diferença significativa para as variáveis idade, gênero, estadia, vulnerabilidade, exames comuns, de imagem recentes e totais. A variável

admissão apresentou $p = 0,006$, tendo o grupo dos demais pacientes obtido valores significativamente maiores que o grupo de discrepâncias (média de $1,85 \times 1,35$ admissões).

A tabela 9 mostra a comparação entre o grupo classificado como vulnerável e os demais, relativo a cada ano e no total e o gráfico 6 mostra a comparação percentual. Pode-se observar que existe um menor percentual de concordância entre os vulneráveis, porém às custas de diminuição das discordâncias menores, sem diferenças significativas entre as maiores.

Tabela 9 – Comparação entre pacientes vulneráveis e não vulneráveis

Ano	Pacientes	Discrep. maiores n (%)	Discrep. Menores n (%)	Concordância n (%)
1980	Vulneráveis	7 (28)	11 (44)	6 (24)
	Não vulneráveis	14 (18,6)	23 (30,6)	35 (46,6)
1990	Vulneráveis	6 (18,1)	16 (48,5)	10 (30,3)
	Não vulneráveis	15 (22,4)	13 (19,4)	38 (56,7)
2000	Vulneráveis	11 (22)	12 (24)	23 (46)
	Não vulneráveis	10 (20)	8 (16)	32 (64)
Total	Vulneráveis	24 (22,2)	39 (36,1)	39 (36,1)
	Não vulneráveis	40 (20,8)	43 (22,3)	105 (54,7)

Gráfico 6 – Comparação de concordância diagnóstica entre pacientes vulneráveis (v) e não vulneráveis (ñ)

Por fim realizamos uma avaliação de concordância inter-observadores com uma amostragem de dez por cento dos casos estudados. Para isto, um outro observador clínico avaliou os prontuários e classificou os casos de acordo com a metodologia de discrepâncias empregada neste estudo. Na categoria de comparação 1 do estudo (diagnóstico clínico final $\times$ necroscópico) obteve-se kappa = 0,67. Na categoria de comparação 2 do estudo (diagnóstico clínico inicial $\times$ necroscópico) obteve-se kappa = 0,51. Na categoria de comparação 3 do estudo (diagnóstico clínico inicial $\times$ clínico final) obteve-se kappa = 0,41. Indicando um poder de concordância moderado (0,41-0,60) e bom (0,61-0,80) entre os observadores para a metodologia utilizada.

Discussão

Caracterização da amostra

A amostra selecionada para o estudo nos diferentes anos não apresentou diferenças significativas de gênero e dias na internação do óbito. O ano de 1980 teve um percentual superior de óbitos selecionados para o estudo em relação aos demais (16,7% dos óbitos $\times$ 7,5% em 1990 $\times$ 6,8% em 2000). Esta diferença de amostragem deveu-se à padronização do estudo em 100 casos por ano em correspondência a estudos semelhantes anteriores (GOLDMAN et al., 1983; SONDEREGGER-ISELI et al., 2000). Outro fator que contribuiu para esta diferença foi o índice de realização de necropsia que foi maior no ano de 1980 (90,9% $\times$ 74,6% em 1990 $\times$ 63,6% em 2000).

A idade dos pacientes selecionados não apresentou diferença importante entre 1980 e 1990. No entanto houve uma diferença importante para o ano de 2000, com uma mediana dos casos cerca de 14 anos maior que a dos outros dois anos e uma média de idade cerca de 9 anos maior. Os pacientes com duas ou mais admissões anteriores no HCFMRPUSP passaram de 12% em 1980 para 25% em 2000.

Das variáveis analisadas para o total de casos a única que apresentou diferença significativa foi a de admissão, com os pacientes com discrepâncias maiores apresentando uma média menor (1,35) que os demais (1,85). Este é um fator importante que pode explicar a melhora na concordância diagnóstica observada na comparação de diagnóstico clínico final $\times$ necroscópico. No entanto, esta melhora não veio com a correspondente diminuição das discrepâncias maiores e sim com uma diminuição das discrepâncias menores. Uma análise

inicial nos leva a correlacionar o incremento no arsenal diagnóstico com a introdução de novos exames e procedimentos com a melhora destes diagnósticos, relativos às discrepâncias menores, visto que não houve diminuição das discrepâncias maiores.

A porcentagem de cadáveres submetidos ao exame necroscópico, apesar de estar em declínio, foi bem elevada em comparação com outros estudos semelhantes na literatura. A manutenção deste percentual elevado difere da maioria dos serviços de patologia que vêm experimentando quedas acentuadas nesta porcentagem ao longo dos últimos anos (BAYER-GARNER; FINK; LAMPS, 2002).

O percentual de realização de necropsias é considerada um fator de influência nos resultados deste tipo de estudo, pois quando apenas uma pequena parte da população falecida é submetida ao exame, possivelmente haverá algum grau de viés na seleção dos casos. Saracci (1991) coloca que, um dos aspectos para a necropsia manter-se como um instrumento de validade para avaliação de diagnósticos clínicos, é haver um adequado procedimento de amostragem para as mortes em estudo.

Causas básicas dos óbitos

A classificação dos casos pela causa básica do óbito mostrou as doenças do aparelho circulatório como as maiores causadoras de morte no grupo selecionado, tendo, no entanto, uma diminuição percentual com o passar dos anos de 43% em 1980 para 30% em 2000. As neoplasias praticamente dobraram sua participação de 12% em 1980 para 23% em 2000. As

doenças infecciosas e parasitárias reduziram à metade os seus casos de 10% em 1980 para 5% em 2000. Este perfil da causa básica de óbito destes casos sinaliza uma tendência do HCFMRPUSP como hospital de referência regional, com um alto número de mortes por neoplasias. A queda nos casos de doenças infecciosas e parasitárias, que passaram de terceira maior causa de óbito em 1980 para a sétima em 2000, parece indicar a mudança do perfil de mortalidade no HCFMRPUSP para o período de tempo do estudo. O alto percentual de mortes por doenças do aparelho circulatório, apesar de sua diminuição, confirma estas doenças como as mais prevalentes em nosso hospital conforme estudos anteriores (KAJIWARA et al., 1993; GRADE, 2003).

A avaliação de precisão, sensibilidade e especificidade, dos maiores grupos diagnósticos de acordo com a classificação da CID 10^a Revisão, mostrou que as doenças do aparelho circulatório tiveram melhora de precisão de 90 para 93%, especificidade de 86 para 93%, mantendo uma sensibilidade intermediária em cerca de 78%. Como este é o grupo com o maior número de casos uma melhora de 3% na precisão é significativa, pois reflete uma melhora em um grupo com um n elevado. As neoplasias mantiveram a precisão em 96% e a especificidade em 95%, obtendo uma melhora na sensibilidade de 75% para 83%. Esta melhora deve-se provavelmente ao aumento de possibilidades de exames diagnósticos no decorrer dos anos do estudo. As doenças do aparelho digestivo mostraram-se com uma baixa sensibilidade que diminuiu ainda mais de 1980 (67%) para 2000 (50%), sendo o terceiro maior grupo, no total dos anos, responsável pela causa básica da morte.

Exame necroscópico completo

Em relação ao exame necroscópico completo que inclui a análise microscópica nossa amostragem mostrou uma tendência de sua menor utilização, indo de 100% dos casos em 1980, para 82% em 1990 e 55% em 2000. Zaitoun e Fernandez (1998) encontraram contribuição significativa para o diagnóstico final, através do exame histopatológico em 5% de sua casuística, além de outros achados patológicos não suspeitados clinicamente em 23% de seus casos. Echenique et al. (2002) afirmaram que o exame histopatológico foi essencial para a correta interpretação dos achados macroscópicos em 17,5% dos casos, em seu estudo. Em nosso estudo a microscopia elucidou o caso, completou ou modificou os diagnósticos macroscópicos maiores em 7% dos pacientes de 1980, 4,9% em 1990 e 7,3% em 2000. Desta forma, pode-se afirmar que a microscopia quando utilizada em nossos pacientes alterou 6,3% dos diagnósticos maiores. Além disso, observamos que por meio do exame histopatológico foi possível detectar doenças clinicamente não suspeitadas (diagnósticos menores) em 20% dos pacientes de 1980, 21,9% em 1990 e 23,6% em 2000. Apesar da diminuição do número de exames necroscópicos que são realizados de maneira completa, não houve diferença nos percentuais de novos achados através da microscopia. Portanto ressalta-se a importância e a necessidade do exame necroscópico completo, para uma maior fidedignidade dos achados patológicos pela necropsia, observado por nossos resultados e também pelos trabalhos da literatura (BERNARDI; SALDIVA; MAUAD, 2005).

Comparação das discrepâncias

Na comparação da categoria 1, entre o diagnóstico clínico final e o necroscópico, nos três anos estudados, pode-se perceber que não houve alteração evolutiva do número de casos enquadrados como discrepâncias maiores (classe I e II). Estes permaneceram em 21 casos a cada ano do estudo, tendo este resultado correlação com alguns estudos anteriores, tanto para análises gerais como para pacientes particularizados, grupos de doenças e ambientes restritos (FERNANDEZ-SEGOVIANO et al., 1988; ROULSON; BENBOW; HASLETON, 2005). Outros estudos demonstraram uma diminuição nas discrepâncias maiores na comparação temporal, sem aumento na concordância (SONDEREGGER-ISELI et al., 2000). Nosso estudo não apresentou melhora das discrepâncias maiores. Houve aumento na concordância diagnóstica de 41 casos em 1980, para 55 em 2000, e estatisticamente significativo com $p = 0,05$. Como não houve melhora nas discrepâncias maiores, fica claro que o aumento da concordância ocorreu pela diminuição das discrepâncias menores (classe III e IV).

Na categoria 2 de comparação, entre o diagnóstico clínico inicial e o necroscópico, pode-se perceber o mesmo padrão de comparação acima descrito. As discrepâncias maiores não apresentaram melhora, mantendo-se sem alterações significativas para seus valores no período do estudo. Houve melhora da concordância diagnóstica de forma estatisticamente significativa com $p = 0,03$. Esta melhora também ocorreu às custas de uma diminuição das discrepâncias menores, de 34 casos em 1980 para 27 em 1990 e 19 em 2000. É interessante notar que os valores absolutos de concordância nesta categoria comparativa são um pouco menores que na categoria anterior (38 x 41 em 1980, 41 x 48 em 1990 e 53 x 55 em 2000), enquanto os valores

absolutos de discrepâncias maiores são mais elevados (24 x 21 em 1980, 30 x 21 em 1990 e 25 x 21 em 2000). Este achado caracteriza-se pelo próprio processo diagnóstico. O diagnóstico clínico final é uma síntese das informações clínicas da entrada dos pacientes, associada aos resultados dos exames realizados e da própria análise da evolução clínica do doente. Desta forma, configura-se um método mais amplo de análise que apresenta maior facilidade de detectar achados discordantes, melhorando, portanto, a concordância.

A melhora da concordância diagnóstica nestas categorias comparativas com a progressão dos anos remete-nos imediatamente à melhora ocorrida nos exames e procedimentos diagnósticos. Estes tiveram um aumento na sua utilização de forma importante, passando de 236 em 1980 para 260 em 1990 e 366 em 2000, tendo aumento na sua utilização maior que 50%. Exames de imagem recentes tiveram aumentos mais significativos como a ultra-sonografia que passou de 2 para 65 exames, a tomografia computadorizada que cresceu de 1 para 59 exames e a ressonância nuclear magnética que passou de 0 para 9 exames.

No entanto, lembramos que a melhora na concordância ocorreu nas duas categorias comparativas, isto é, tanto na categoria de diagnóstico clínico inicial como no clínico final comparando-se com o diagnóstico necroscópico. Como o inicial foi coletado sem o uso dos exames e procedimentos diagnósticos discriminados no estudo, devemos discutir outras possibilidades para a tendência verificada.

Dos dados coletados, podemos observar que houve um aumento do número de pacientes com duas ou mais internações prévias no próprio serviço de 12 para 25% dos casos estudados. Houve também uma diminuição do número de pacientes sem admissões anteriores de 64 para 60%. Esta tendência sinaliza que o conhecimento prévio destes doentes e de seus possíveis diagnósticos ajudou na melhora da concordância.

Outro fator que pode ter contribuído para esta melhora é o encaminhamento médico. Como o HCFMRPUSP é um hospital de referência regional, muitos dos casos encaminhados já fizeram uma avaliação médica inicial. Este argumento valeria para os três anos do estudo. No entanto, a estrutura de saúde do município e da região vem aumentando progressivamente, proporcionando um maior acesso aos serviços de saúde para a população de uma maneira geral. Além disso, em 1999 estava iniciando o processo de hierarquização dos serviços de saúde na cidade e região com a implementação da Central de Regulação Médica. Isto significa que a maioria dos pacientes internados em 2000 passou por uma avaliação médica inicial antes de sua internação.

Na categoria três de comparação entre os passos diagnósticos, clínico inicial e final, observa-se claramente o alto percentual de concordância. Se analisarmos o processo diagnóstico com a avaliação inicial constituindo a anamnese, associada ao exame físico com elaboração de hipóteses diagnósticas que geralmente são testadas com exames complementares, observamos que este alto percentual de concordância é realmente o esperado. Ele mantém-se no patamar de 80% em todos os anos do estudo. Por outro lado, também pode significar uma dificuldade de realização de um número maior de hipóteses diagnósticas, assim como da reavaliação constante dos diagnósticos inicialmente efetuados. Outro dado que podemos observar é uma tendência de aumento nas discrepâncias diagnósticas classe I, que foram de 3 em 1980, 6 em 1990 e 9 em 2000. À análise estatística a comparação ao longo das décadas apresentou $p = 0,07$, que não corresponde a significância estatística, porém está muito próximo do nível de significância, demonstrando forte tendência de crescimento. Quando analisamos as discrepâncias maiores (classe I e II) não foi possível verificar tal tendência.

Ao analisarmos a comparação entre o diagnóstico clínico inicial e o final e observarmos que a concordância mantém-se em um nível muito elevado de 80%, tendemos a considerar um dado esperado e sem utilidade. Os diagnósticos clínicos iniciais quando suspeitados dificilmente apresentam mudanças ao longo do processo diagnóstico. Avaliamos, então, o diagnóstico clínico inicial dos pacientes que na categoria comparativa 1 (clínico final $\times$ necroscópico) apresentaram discrepâncias classe I e observamos que em 89% deles havia uma concordância entre o diagnóstico clínico inicial e o final. Portanto com um percentual de concordância superior aos demais pacientes.

Com estes dados reforça-se um possível tipo de erro formulado por Voytovich, Rippey e Suffredini (1985) denominado de encerramento prematuro do diagnóstico e explicitado por Sonderegger-Iseli et al. (2000) em seu sistema para erros. Este resultado mostra que o encerramento prematuro dos diagnósticos foi um dos principais mecanismos responsáveis pelas discrepâncias em nosso estudo.

Exames complementares

Em relação aos exames e procedimentos diagnósticos que foram pesquisados em nosso estudo houve um aumento importante, como descrito acima, do número de exames totais. Estes passaram de 236 em 1980 para 366 em 2000. Em 1980 e 1990 seis pacientes em cada ano não realizaram nenhum dos exames pesquisados, enquanto que em 2000 apenas dois pacientes não foram submetidos a nenhum destes procedimentos diagnósticos. O percentual de concordância

do laudo destes exames com os achados de necropsia foi de 95% em 1980, 93% em 1990 e 98% em 2000. Estes números mostram uma tendência de melhora da concordância que sempre esteve em níveis elevados.

Os exames radiológicos não contrastados apresentaram discreta diminuição no seu uso de 139 em 1980 para 117 em 2000 com uma queda de 15%, mantendo, no entanto, uma alta utilização de cerca de 1,2 radiografias por paciente internado em todo o estudo. O percentual de concordância do laudo com os achados de necropsia ficaram em torno de 95% nos três anos estudados. Por serem não invasivos, altamente informativos e com efeitos colaterais mínimos mantêm-se como os exames mais solicitados dentre todos os estudados.

Os exames radiológicos contrastados tiveram uma diminuição importante, passando de 39 em 1980 para 14 em 2000. A queda na sua utilização foi de 64%. Em 1980 era o segundo exame mais solicitado, sendo o quarto em 1990 e o sétimo em 2000. O percentual de concordância do laudo com os achados de necropsia foi de 90% em 1980 e 100% em 1990 e 2000. Estes exames estão com seu uso em declínio por fatores como o advento de outros exames mais modernos, com melhor definição de imagem e por necessitar do uso de substâncias contrastantes com possíveis efeitos colaterais, como reação alérgica a contrastes.

As cintilografias mantiveram uma taxa de utilização em cerca de 10% dos pacientes nos anos estudados, apresentando um percentual de concordância de 67% em 1990 e 100% em 1980 e 2000. Este é um exame que não teve alterações importantes e que mantém indicações precisas para vários diagnósticos.

As explorações cirúrgicas apresentaram uma queda de 62,5% na sua utilização, sendo utilizadas em 8 vezes em 1980, 6 em 1990 e 3 em 2000. A operação cirúrgica como ferramenta diagnóstica manteve um percentual de 100% de concordância com os achados de necropsia. O

declínio em seu uso está relacionado com a aquisição de novas tecnologias que permitem um diagnóstico mais preciso sem a necessidade de procedimentos tão agressivos, além da melhora constante da terapêutica que permite diminuir a exposição de pacientes a riscos cirúrgicos.

As endoscopias tiveram um aumento importante passando de 13 exames em 1980 para 32 em 2000, com um aumento de 146% na sua utilização. Estes exames são invasivos, porém utilizam-se de orifícios naturais do corpo humano, conseguindo acesso ao interior do mesmo sem expor o paciente a grandes riscos. Com a melhora da tecnologia de micro-câmeras estes exames passaram a ser menos desconfortáveis e mais seguros, com conseqüente aumento de sua utilização.

As biópsias também mostraram um aumento importante e semelhante ao das endoscopias de 147% nos três anos do estudo, passando de 23 em 1980 para 43 em 1990 e 57 em 2000. É interessante notar que mesmo sendo exames invasivos apresentaram uma tendência de aumento, diferentemente dos procedimentos cirúrgicos. Por serem menos invasivos que as explorações cirúrgicas, diminuindo os riscos destas, houve esta tendência. O percentual de concordância com os achados de necropsia foi alto em 1980 com 96% e 2000 com 98%, tendo, no entanto, apresentado um valor baixo em 1990 que foi de 88%. Não formamos uma compreensão sobre este dado.

Classificamos como exames de imagem recentes os exames de imagem de ultrassonografia, tomografia computadorizada e ressonância nuclear magnética. Esta divisão deveu-se também ao grande aumento de sua utilização.

As ultra-sonografias foram utilizadas duas vezes em 1980, 35 em 1990 e 65 vezes em 2000, tendo, portanto, um aumento de 3150% no seu uso. Em 1980 tiveram um percentual de concordância com os achados de necropsia de 50%, que evoluiu para 94% em 1990 e 98% em

2000. Este perfil de concordância corresponde ao próprio desenvolvimento do exame e de seus operadores. Por ser um exame dinâmico e com uma interpretação muito dependente do operador observamos que com o passar dos anos houve uma melhora da concordância, que deve ser relacionada à sua melhor interpretação e ao melhor treinamento dos ultra-sonografistas. É um exame não invasivo que oferece inúmeras informações internas do paciente, e principalmente por estas características, apresentou este incremento no seu uso.

As tomografias computadorizadas tiveram um aumento de 5800% na sua utilização no período estudado. Em 1980 apenas um paciente realizou uma tomografia, que foi feita fora do hospital em outro serviço, pois o HCFMRPUSP não dispunha deste exame. Em 1990 foram 14 exames e em 2000 foram 59. O percentual de concordância do laudo com os achados de necropsia foi de 100% em todo o período do estudo. Por ser um exame com demonstração de imagens radiográficas reconstruídas em diversos planos de corte e em vários ângulos, ele possibilita uma análise apurada das mais diversas estruturas do corpo humano. Esta característica associada ao fato de ser um exame não invasivo fez que sua utilização trouxesse grandes informações para a prática médica geral. Foi o exame que apresentou o maior percentual de aumento de utilização seguido pela ultra-sonografia.

A ressonância nuclear magnética começou a ser utilizada apenas no meio da década de 90. Com isto, em 1980 e 1990 não houve paciente submetido a este exame. Em 2000 ela foi usada 9 vezes nos paciente avaliados. Apresentou uma concordância de 100% de seu laudo com os achados de necropsia. É um exame com características semelhantes às descritas para a tomografia computadorizada, de poder de análise e capacidade de obtenção de informações. Além disto não utiliza radiações para a sua realização, tornando-se mais segura, e fornece imagens mais nítidas que a tomografia permitindo estudos de estruturas mais delicadas.

Discrepâncias maiores

Ao analisarmos as discrepâncias de classe II, que foram descritas na tabela 6, discriminadas por ano e sistema que a apresentou, podemos perceber que os diagnósticos encontrados na necropsia e que não foram realizados durante a estadia do doente em nosso hospital são variados e pouco se repetem. Em 1980 a maior parte destes diagnósticos foram relacionados com o sistema cardiovascular, sendo que o único que se repetiu foi a broncopneumonia em dois casos de um total de dez. Em 1990 novamente a broncopneumonia apresentou repetição, desta vez de três casos em um total de oito. Em 2000 os quadros ligados ao sistema respiratório predominaram com cinco casos, mas o diagnóstico que não foi realizado e que se repetiu duas vezes foi o de acidente vascular cerebral. No total, os diagnósticos enquadrados nesta categoria que mais aconteceram foram as broncopneumonias com cinco casos (19,3% do total) e os acidentes vasculares cerebrais com quatro casos (15,4% do total). A broncopneumonia é uma das causas mais comuns de internações hospitalares e quando não relacionada com a causa da internação costuma ser responsável por agravamento de quadros em pacientes com internações prolongadas. O acidente vascular cerebral pode ser de difícil diagnóstico em pacientes com dificuldade de comunicação e mesmo em idosos comunicantes, quando não são extensos e com localizações particulares. A tendência que se poderia observar com estes dados seria a do melhor diagnóstico de broncopneumonia em 2000, visto que neste ano, este diagnóstico não ocorreu na necropsia de nenhum dos pacientes do estudo. No entanto, ao confrontar este diagnóstico na tabela 7, observamos que em 2000 por duas vezes este diagnóstico esteve relacionado com o óbito de forma direta e que se fosse realizado

anteriormente poderia ter mudado o prognóstico do doente. Portanto é um diagnóstico que por ser muito prevalente deve ser pesquisado com frequência nos pacientes internados.

Ao observarmos a tabela 7 com a descrição e distribuição dos diagnósticos discrepantes classe I discriminados por cada ano de estudo, pode-se perceber em 1980 a peritonite como o diagnóstico mais freqüente nesta categoria, acontecendo por três vezes em um total de onze casos. As doenças relacionadas com o sistema gastrointestinal prevaleceram no total dos anos e em 1980 e 2000 isoladamente. Em 1990 aconteceram três casos não diagnosticados de edema agudo de pulmão com insuficiência cardíaca em que não houve o início de terapêutica para edema agudo, mesmo com quadro de dispnéia. Aconteceram também três casos de úlceras pépticas perfuradas não diagnosticadas, tanto em 1990 como em 2000. Os demais diagnósticos não se repetiram a não ser a broncopneumonia como discutido acima. As úlceras aconteceram em sete pacientes (18,9% do total) seguidas de quatro casos de insuficiência cardíaca com edema agudo pulmonar (10,8% do total), três casos de peritonite (8,1% do total) e três casos de broncopneumonia.

O achado da falta de diagnóstico de úlceras pépticas perfuradas com hemorragia digestiva alta ou peritonites devido a outras causas foi descrito na literatura anteriormente (TSE; LEE, 2003). Este diagnóstico pode ser difícil de realizar em pacientes especiais ou com dificuldade de comunicação e deve ser lembrado para pacientes submetidos a internações prolongadas com alto grau de estresse. O edema agudo de pulmão é um diagnóstico clínico e apóia-se na observação e exame físico do doente, portanto um acompanhamento próximo de pacientes com diversos riscos associados é importante para não deixar passar este agravo. Em 2000 não houve nenhum caso de edema agudo de pulmão nas discrepâncias maiores.

Na tentativa de aprofundar a análise das discrepâncias classe I em relação aos exames e procedimentos diagnósticos estudados, nós avaliamos a adequação destes exames em cada caso. Por meio da tabela 8 podemos observar que no decorrer dos anos do estudo o percentual de exames que poderiam sugerir o diagnóstico e não foram requisitados permaneceu ao redor de 55% dos casos. Os casos que não conseguiram ser classificados diminuíram de 45% em 1980, para 23% em 1990 e 15% em 2000. Associado a esta tendência observamos um aumento de exames apropriados para o diagnóstico que possivelmente tiveram uma interpretação equivocada que foi zero em 1980, 15% em 1990 e 30% em 2000.

Estes números nos remetem a uma nova situação que pode ocorrer com o incremento no número de exames e procedimentos diagnósticos ao longo dos anos, isto é, a interpretação equivocada dos novos exames. Esta tendência deve diminuir com o uso mais disseminado de novas tecnologias e com a melhora no conhecimento médico acerca da interpretação dos novos exames em associação com os achados de exame físico. No entanto, é interessante notar que eles demonstram que deve haver um cuidado com as novas tecnologias, e uma incorporação gradual e crítica de todas elas.

Concordância inter-observadores

No intuito de fazer uma avaliação do nível de subjetividade da classificação adotada, optamos por verificar o grau de concordância entre diferentes observadores, ao avaliar os mesmos casos. Para isto foi feita uma nova amostragem de 10% dos casos avaliados para que

um terceiro observador fizesse a classificação. Os resultados mostram que o melhor resultado de concordância ($\kappa = 0,67$) foi na comparação entre o diagnóstico clínico final e o necroscópico. Este resultado é de certa forma esperado, visto que nesta categoria os diagnósticos estão expressos com maior clareza nos prontuários e sua interpretação passa a ser menos subjetiva. Quando a comparação envolveu o diagnóstico clínico inicial, o nível de concordância apresentou uma diminuição ($\kappa = 0,51$ x necroscópico e $\kappa = 0,41$ x clínico final). Estatisticamente, são consideradas concordâncias moderadas, mas conseguem demonstrar uma maior dificuldade de classificação dos casos quando colocamos a nova variável introduzida por este trabalho que é a do diagnóstico clínico inicial.

Apesar destas classificações adotadas apresentarem este grau de subjetividade, sendo portanto um possível viés do trabalho, ao realizarmos esta comparação inter-observadores pudemos mostrar que ela pode ser realizada com um grau de concordância de moderada a boa. Outro procedimento que procuramos realizar para diminuir este grau de subjetividade foi de colocar dois pesquisadores, um clínico e um patologista, para realizar a classificação em conjunto. Esta troca de experiências não ocorreu com o observador que foi usado para avaliar a concordância, visto que trabalhou sozinho.

Pacientes vulneráveis

A variável criada para uma outra abordagem à comparação das discrepâncias diagnósticas que nomeamos de vulneráveis, propiciou uma avaliação de possíveis negligências

em relação a grupos especiais de pacientes. O conceito de vulnerabilidade foi incorporado aos debates bioéticos nos últimos anos, mais especificamente na década de 90. Historicamente, subentendeu-se que os vulneráveis seriam os deficientes mentais, deficientes físicos, as crianças, os senis e os institucionalizados de qualquer ordem, mas foram deixadas de lado todas as pessoas que se encontravam em situação de vulnerabilidade. Vulnerável é toda pessoa que se encontra menos apta a se proteger (DINIZ; CORREA, 2001). Lau et al (2005) em um trabalho com idosos, discute os fatores relacionados com uma maior potencialidade de prescrição médica inadequada para este grupo de pessoas. Optamos por agrupar além dos idosos, os pacientes etilistas crônicos, ou com seqüelas neurológicas ou com malformações congênitas do sistema nervoso. Desta forma, não incluímos as crianças e os deficientes físicos por considerar que estes pacientes não apresentam situação de vulnerabilidade no seu atendimento médico em um hospital, com a equipe de saúde dispendendo muitas vezes maiores cuidados com estes pacientes do que com os demais. Os pacientes institucionalizados de qualquer ordem não poderiam ser distinguidos com as informações constantes nos prontuários e, portanto, optamos por não colocá-los neste grupo. Os etilistas crônicos, pacientes com seqüelas neurológicas ou com malformações congênitas do sistema nervoso central foram escolhidos por considerarmos que estão em situação de vulnerabilidade no atendimento médico em um hospital geral. Portanto, esta escolha deveu-se à observação de que estes pacientes são muitas vezes atendidos sem a retaguarda de melhores cuidados, por apresentarem estas condições especiais.

O número de pacientes que se encaixavam no grupo de vulneráveis passou de 24 em 1980, para 32 em 1990 e 46 em 2000. Houve, portanto, um aumento importante de pacientes com estas características durante o período de tempo do estudo, chegando a quase metade da amostra de 2000. Este dado demonstra uma alteração do perfil dos pacientes internados pelo

HCFMRPUSP, sendo coerente com o aumento da média de idade dos pacientes verificadas em 2000 em relação aos outros anos.

Em 1980 a diferença entre as discrepâncias maiores (classe I e II) encontradas nos pacientes vulneráveis ficou evidente com 28% x 18,6% nos demais. As discrepâncias menores também estiveram mais altas neste grupo de pacientes com 44% x 30,6%. No entanto a maior diferença encontra-se no percentual de concordâncias entre estes dois grupos de pacientes, sendo de 46,6% para os não vulneráveis e 24% para os vulneráveis.

As discrepâncias maiores não apresentaram diferença importante entre estes dois grupos comparativos em 1990 e 2000, inclusive em 1990 houve um número menor destas entre os vulneráveis (18,1% x 22,4%). As discrepâncias menores foram maiores nestes pacientes também em 1990 e 2000, tendo uma diferença maior entre os grupos de 1990 (48,5% x 19,4%). As concordâncias em 1990 e 2000 mantiveram o mesmo perfil de 1980 com um valor menor de concordância para o grupo de pacientes vulneráveis. É interessante notar que o percentual de concordâncias diagnósticas apresentou melhora no período de tempo do estudo, tanto no grupo de vulneráveis (24% em 1980, 30,3% em 1990 e 46% em 2000) quanto no outro grupo (46,6% em 1980, 56,7% em 1990 e 64% em 2000). Entretanto demonstra-se que existe uma diferença importante entre estes dois grupos de pacientes.

O fato do grupo de pacientes classificados como vulneráveis ser composto por pessoas com necessidades especiais que comumente apresentam várias doenças associadas é um fator importante para a presença da diferença de concordância entre os grupos. Em pacientes que têm um número menor de doenças temos uma maior chance de concordância de todos os diagnósticos, os maiores e os menores. Neste sentido, nota-se que para as discrepâncias maiores houve nos anos de 1990 e 2000 um número semelhante para os dois grupos, mostrando uma

maior preocupação dos cuidadores com os principais diagnósticos destes doentes, aqueles que colocam em risco a vida do paciente. Outro fator que contribui para estas diferenças é a dificuldade de comunicação que estes doentes podem apresentar. A falta de comunicação adequada impede que o médico que está cuidando deles fique a par de todas as doenças progressas e de todos os sintomas que estes apresentam.

Portanto, nota-se em cada ano do estudo e no total que os pacientes classificados como vulneráveis tiveram um percentual de concordância menor que os pacientes não vulneráveis. Infelizmente, as características de serem pacientes com múltiplas doenças e com complicações maiores que os outros doentes podem não ser os únicos fatores responsáveis por esta diferença. O motivo por termos selecionado estes pacientes, e nomeado este grupo de vulneráveis, isto é, por serem doentes com necessidades especiais, pode ser outro fator importante para os resultados de concordância diagnóstica encontrados sugerindo que eles possam ser tratados de forma diferenciada em relação aos pacientes considerados “normais”.

Considerações Finais

Na comparação de todos os pacientes, entre o diagnóstico clínico final e o necroscópico, as discrepâncias maiores não apresentaram melhora na avaliação evolutiva. Este resultado corrobora outros estudos prévios, que analisaram várias populações, grupos de doenças, pacientes de diferentes faixas etárias, ou ambientes hospitalares particularizados (BRITTON, 1974; BURTON; UNDERWOOD, 2007; CAMERON, 1984; FERNANDEZ-SEGOVIANO et

al., 1988; ROULSON; BENBOW; HASLETON, 2005; XAVIER et al, 2005). Apesar do aumento do número de exames e procedimentos diagnósticos, poucos estudos têm demonstrado diminuição nas discrepâncias diagnósticas maiores (SHOJANIA et al., 2003; SONDEREGGER-ISELI et al., 2000; VADILLO et al., 2006). Nossa avaliação não encontrou diferença estatística entre os grupos de pacientes com concordância diagnóstica e de discrepâncias maiores, quanto ao número de exames e procedimentos realizados. Estes achados estão de acordo com vários autores, que inclusive discutem a necessidade de cuidado no excesso de confiança nos exames modernos, pois eles podem eventualmente contribuir diretamente para o acontecimento das discrepâncias maiores (GOLDMAN, 1984; KIRCH; SCHAFFI, 1996).

Nosso estudo apresentou um aumento significativo na concordância entre os diagnósticos clínicos iniciais e finais em relação aos diagnósticos necroscópicos. Este aumento da concordância foi concomitante à diminuição das discrepâncias menores, sem alteração percentual, na avaliação entre as décadas, das discrepâncias maiores. A redução das menores deve estar relacionada com o aumento no número de realização de exames e procedimentos diagnósticos, assim como com o aumento do número de pacientes com duas ou mais internações prévias no próprio hospital.

Este aumento das concordâncias diagnósticas entre as décadas estudadas esteve presente nos dois grupos separados para avaliação, os vulneráveis e os não vulneráveis. No entanto houve uma clara diferença no percentual de concordâncias entre estes dois grupos, com os vulneráveis apresentando um número menor em relação ao outro grupo. Esta diferença caracteriza-se nas discrepâncias menores, visto que as maiores apresentaram valores semelhantes para os dois grupos, à exceção do ano de 1980. Estes resultados demonstram que

existe uma preocupação quanto aos diagnósticos relacionados com a causa básica ou direta do óbito para os dois grupos.

De maneira esperada, a concordância entre o diagnóstico clínico inicial e o final permaneceu em 80% nos três períodos estudados. Interessante, para os casos em que houve discrepância classe I na comparação entre o diagnóstico clínico final $\times$ necroscópico, a concordância entre o diagnóstico clínico inicial $\times$ final foi de 89%. Estes números sugerem que existe uma resistência importante na mudança das hipóteses diagnósticas iniciais por parte dos médicos.

Graber et al. estudaram cem erros diagnósticos envolvendo médicos, vários deles sendo descobertos pela análise necroscópica, e tiveram como resultados que os erros cognitivos mais freqüentes foram os relacionados com o encerramento prematuro do diagnóstico. Isto é, a falta de continuar considerando alternativas razoáveis após realizar o diagnóstico inicial (GRABER; FRANKLIN; GORDON, 2005).

Sonderegger-Iseli et al. (2000) propuseram um sistema de avaliação de erros que ajudam a avaliar faltas cognitivas. Neste sistema o encerramento prematuro do diagnóstico é o mecanismo mais comum de levar a erros, pois este mecanismo pode estar presente na geração da hipótese, nas fases de história e exame físico, desencadeamento (“triggering”) ou de formulação contextual. Pode ainda estar presente como mecanismo de erro na coleta e processamento de informações e na verificação destas.

Portanto, a melhora na qualidade do diagnóstico inicial é crucial para a melhora da precisão do processo diagnóstico. Outro aspecto importante é a educação permanente do médico para que ele esteja sempre atento no processamento das informações e exames e na verificação do diagnóstico, continuando a considerar alternativas razoáveis ao primeiro diagnóstico.

Vadillo et al. (2006) apresentaram o único trabalho prospectivo que obteve uma redução nas discrepâncias diagnósticas maiores após um processo de intervenção. Esta foi a introdução de uma análise sistemática dos diagnósticos anteriores à necropsia realizados pelos clínicos que atenderam os pacientes, confrontados com os diagnósticos necroscópicos. Esta sistematização, com retro-alimentação contínua entre clínicos e patologistas, demonstrou capacidade de diminuir as discrepâncias.

Nossos achados colocam a necessidade de uma discussão mais abrangente sobre o real papel da moderna tecnologia no processo de elaboração do diagnóstico. Com os inúmeros avanços nos exames e procedimentos diagnósticos, pode-se observar que tem havido um aumento natural e esperado da crença na ciência e tecnologia. Por outro lado, esta confiança em análises muitas vezes quantitativas conduzidas em um laboratório, sobrepõe-se a uma análise qualitativa conduzida subjetivamente pelo médico (REISER, 1978). Este aspecto está em acordo com a visão de que, durante o século 20, ocorreu uma mudança na qual nossa confiança na tecnologia transformou-a de um meio para um fim. Muitas vezes o desenvolvimento tecnológico não está nascendo a partir de uma necessidade de resolução de um problema, mas, ao contrário, como uma manifestação baseada na crença de que se esta nova tecnologia pode ser inventada, ela deve ser (POSTMAN, 1992). Esta mudança, em nossa avaliação, vem levando o médico moderno a deixar de investir no contínuo desenvolvimento da arte do diagnóstico, considerada por muitos como imprecisa e demorada.

Sanders (2005) coloca apropriadamente a tecnologia médica em perspectiva, quando propõe uma resposta simples, afirmativamente ou negativamente, para a questão se a tecnologia têm nos levado a realizar uma medicina melhor. A resposta é afirmativa se

pudermos observar apenas pela ótica das incríveis e maravilhosas oportunidades que podem ser realizadas com a tecnologia, como os novos tipos de procedimentos, a minoração da dor, novos tipos de procedimentos cirúrgicos, coleta de dados mais precisos, entre outros. Por outro lado, a resposta é negativa se considerarmos que muitos procedimentos desnecessários ou que têm potencialidade para levar a outros problemas são realizados porque eles simplesmente podem ser realizados. Sanders (2005) considera que “a tecnologia pode não apenas distanciar o médico do paciente, mas também o distanciar dele mesmo, de sua história, e da história de sua profissão”, citando também Weiss RJ da Universidade de Columbia que considera que “para uma profissão que se considera científica, nós temos abordado o assunto do uso apropriado da tecnologia diagnóstica e de tratamento em um modo não-científico. Não apenas temos sido não-científicos, não temos sido humanitários”.

O fato de diversos estudos apresentarem o quadro de manutenção das discrepâncias diagnósticas que nossos dados corroboram, também deve ser usado para abordar outro assunto pouco discutido na literatura médica que é de uma “necessária falibilidade” dos médicos (ANDERSON; HILL; KEY, 1989). Mesmo aceitando que deve haver a busca contínua de uma prática médica sem erros ou discrepâncias, é necessário que haja uma discussão acerca da falibilidade natural de qualquer processo. A elaboração de um diagnóstico é um processo com inúmeras variáveis, e, portanto, sempre sujeito a vários tipos de falhas.

O médico, como qualquer outro profissional que faz parte de um sistema, é falível e têm uma grande dificuldade de lidar com esta verdade (KAAFARANI; ITANI, 2005). A crença irracional na tecnologia pode nos levar a esquecer estes aspectos, principalmente da complexidade que é o processo diagnóstico, e isto pode ter conseqüências importantes em questões clínicas, econômicas, éticas e forenses.

O ensino, a pesquisa e a extensão que são a base da atividade acadêmica, constituem áreas em que pode haver uma contribuição para a redução de erros médicos e a melhora da segurança do paciente. Com o objetivo de melhora na qualidade do atendimento médico e na segurança do paciente dentro do sistema de saúde, várias medidas estão sendo propostas e realizadas nos países desenvolvidos (BARACH; SMALL, 2000; INSTITUTE OF MEDICINE, 1999; MEYER, 2001). O estudo e o debate sobre este assunto na literatura científica internacional estão cada vez mais presentes. É importante que possamos desenvolver estudos e propostas neste sentido em nossas instituições para que comecemos a tratar esta questão de maneira mais científica e menos emocional.

Conclusões

Não observamos melhora das discrepâncias diagnósticas maiores no período abrangido pelo estudo nas três categorias de avaliação pesquisadas. Houve melhora estatisticamente significativa da concordância diagnóstica nas comparações entre os diagnósticos clínicos (inicial e final) e o necroscópico. Esta melhora baseou-se apenas na diminuição das discrepâncias menores. A comparação entre o diagnóstico clínico inicial e final mostrou elevada concordância diagnóstica ao redor de 80%.

O exame necroscópico completo mostrou uma tendência de menor utilização, de 100% dos casos em 1980 para 55% em 2000. A microscopia, contudo, elucidou o caso, completou ou modificou os diagnósticos macroscópicos maiores em 6,3% destes casos e mostrou doenças clinicamente não suspeitadas (diagnóstico menores) em 21,5%.

Os diagnósticos das doenças do aparelho circulatório apresentaram melhora de precisão (“*accuracy*”) de 90 para 93%, de especificidade de 86 para 93%. As neoplasias mantiveram a precisão em 96% e a especificidade em 95%, obtendo uma melhora na sensibilidade de 75 para 83%. As doenças do aparelho digestivo mostraram-se com uma baixa sensibilidade que diminuiu ainda mais de 1980 (67%) para 2000 (50%), sendo o terceiro maior grupo, no total dos anos, responsável pela causa básica da morte.

O grupo classificado como vulnerável apresentou menor concordância diagnóstica em relação aos demais em todos os anos do estudo. Houve uma melhora da concordância para os dois grupos na avaliação temporal com melhores percentuais para os não vulneráveis (46,6% e 24% em 1980, 56,7% e 30,3% em 1990, 64% e 46% em 2000).

Nos pacientes com discrepâncias maiores de classe I, observamos que os exames que poderiam contribuir para o estabelecimento do diagnóstico não foram requisitados em 55% dos

casos, ou quando realizados, possivelmente apresentaram uma interpretação equivocada em 16,2% deles.

Nossos achados enfatizam a necessidade de uma discussão mais abrangente sobre o real papel da moderna tecnologia no processo de elaboração do diagnóstico. Além disso, mesmo aceitando que deve haver a busca contínua de uma prática médica sem erros ou discrepâncias, é necessário que haja uma discussão acerca da falibilidade natural deste processo, pois apresenta inúmeras variáveis, e, portanto, sempre sujeito a vários tipos de falhas. O médico, como qualquer outro profissional que faz parte de um sistema, é falível. A crença irracional na tecnologia pode nos levar a esquecer estes aspectos, e isto pode ter consequências importantes em questões clínicas, econômicas, éticas e forenses.

Referências Bibliográficas*

* O modelo de referência está de acordo com o *International Committee of Medical Journal Editors* (Vancouver Style) – Grupo de Vancouver

- Aalten CM, Samson MM, Jansen PA. Diagnostic errors; the need to have autopsies. *Neth J Med* 2006;64(6):186-190.
- Altman DG. Practical statistics for medical research. London: Chapman and Hall; 1991.
- American Medical Association - Council on scientific affairs. Autopsy - a comprehensive review of current issues. *JAMA* 1987;258(3):364-369.
- Anderson NH, Shanks JH, McCluggage GW, Tonor PG. Necropsies in clinical audit. *J Clin Pathol* 1989;42:897-901.
- Anderson RE, Hill RB. The autopsy in academic medical centers in the United States. *Hum Pathol* 1988;19:1369-1371.
- Anderson RE, Hill RB, Key CR. The sensitivity and specificity of clinical diagnostics during five decades – toward an understanding of necessary fallibility. *JAMA* 1989;261:1610-1617.
- Barach P, Small SD. Reporting and preventing medical mishaps: systems lessons from non-medical near miss reporting. *BMJ* 2000;320:759-763.
- Barber N. Designing information technology to support prescribing decision making. *Qual Saf Health Care* 2004;13:450–454.
- Baron JH. Clinical diagnosis and the function of necropsy. *J R Soc Med* 2000; 93 : 463-6.
- Bates DW. Using information technology to reduce rates of medication errors in hospitals. *BMJ* 2000;320:788-791.
- Battle RM, Pathak D, Humble CG, Key CR, Vanatta PR, Hill RB, et al. Factors influencing discrepancies between premortem and postmortem diagnoses. *JAMA* 1987;258:339-344.
- Bayer-Garner IB, Fink LM, Lamps LW. Pathologists in a teaching institution asses the value of the autopsy. *Arch Pathol Lab Med* 2002;126:442-447.

- Belzeberg H, Rivkind AI. It always pays to make a diagnosis: to autopsy or not to autopsy? *Crit Care Med* 1998;26:1299-1300.
- Bernardi FDC, Saldiva PHN, Mauad T. Histological examination has a major impact on macroscopic necropsy diagnoses. *J Clin Pathol* 2005;58:1261-1264.
- Bierig JR. A potpourri of legal issues relating to the autopsy. *Arch Pathol Lab Med* 1996;120:759-762.
- Blosser AS, Zimmerman HE, Stauffer JL. Do autopsies of critically ill patients reveal important findings that were clinically undetected? *Crit Care Med* 1998;26(8):1332-1336.
- Bobb A, Gleason K, Husch M, Feinglass J, Yarnold PR, Noskin GA. The epidemiology of prescribing errors. The potential impact of computerized prescriber order entry. *Arch Intern Med* 2004;164:785-792.
- Borrel-Carrió F, Epstein RM. Preventing errors in clinical practice: a call for self-awareness. *Ann Fam Med* 2004;2:310-316.
- Böttiger LE. The post-mortem: its decline and fall [Editorial]? *J Intern Med* 1992;231:99-101.
- Brennan TA, Leape LL, Laird NM, Hebert L, Localio AR, Lawthers AG, et al. Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard medical practice study I. *N Engl J Med* 1991;324(6):370-376.
- Britton M. Clinical diagnostics: experience from 383 autopsied cases. *Acta Med Scand* 1974;196:211-219.
- Buckner T, Blatt J, Smith SV. The autopsy in pediatrics and pediatric oncology: a single-institution experience. *Pediatr Dev Pathol* 2006;9(5):374-380.
- Burke CS, Salas E, Wilson-Donnelly K, Piest H. How to turn a team of experts into an expert medical team: guidance from the aviation and military communities. *Qual. Saf. Health Care* 2004;13:96-104.

- Burton EC, Troxclair DA, Newman WP. Autopsy diagnoses of malignant neoplasms – how often are clinical diagnoses incorrect? JAMA 1998;280:1245-1248.
- Burton JL, Underwood J. Clinical, educational, and epidemiological value of autopsy. Lancet 2007;369:1471-1480.
- Cabot RC. Diagnostic pitfalls identified during a study of 3000 autopsies. JAMA 1912;59:2295-2298.
- Cameron HM. The autopsy, past and present. J R Coll Physicians Lond 1984;18(4):236-239.
- Cameron HM, McGoogan E. A prospective study of 1152 hospital autopsies: I - Inaccuracies in death certification. J Pathol 1981;133(4):273-283.
- Chernof D. The role of managed health care organizations in autopsy reimbursement. Arch Pathol Lab Med 1996;120(8):771-772.
- Chute CG, Ballard DJ, Nemetz PN. Contribution of autopsy to population - based cancer epidemiology: target intervention to improve ascertainment. IARC Sci Publ 1991;112:207-216.
- Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas 1960;20:37-39.
- Combes A, Mokhtari M, Couvelard A, Trouillet JL, Baudot J, Henin D, et al. Clinical and autopsy diagnoses in the intensive care unit: a prospective study. Arch Intern Med 2004;164(4):389-392.
- Cordner SM. Deciding the cause of death after necropsy. Lancet 1993;341:1458-1460.
- Denison A, Pierce JR. Systems analysis of a clinical error. J Public Health Manag Pract 2003;9(1),43–46.
- Deskin WC, Hoye RE. Another look at medical error. J Surg Oncol 2004;88:122-129.

- Dhar V, Perlman M, Vilela MI, Haque KN, Kirpalani H, Cutz E. Autopsy in a neonatal intensive care unit: utilization patterns and associations of clinicopathologic discordances. *J Pediatr* 1998;132:75-79.
- Diamond I. New approach needed to revive autopsy. *Arch Pathol Lab Med* 1996;120:713.
- Diniz D, Corrêa M. Declaração de Helsinque: relativismo e vulnerabilidade. *Cad Saúde Pública* 2001;17(3):679-688.
- Dunn KL, Reddy P, Bowes G. Uses of error: diagnosis, detection, and disclosure. *Lancet* 2005;365:386.
- Echenique LS, Mello RA, Odashiro LN, Franco M. Correlação entre achados macro e microscópicos em 200 autópsias consecutivas: análise do valor custo/benefício do estudo histopatológico completo das autópsias. *Jornal Brasileiro de Patologia e Medicina Laboratorial* 2002;38(3):219-224.
- Egerváry M, Szende B, Roe FJ, Lee PN. Accuracy of clinical diagnosis of lung cancer in Budapest in an institute specializing in chest diseases. *Pathol Res Pract* 2000;196(11):761-766.
- Ermenc B. Discrepancies between clinical and post-mortem diagnoses of causes of death. *Med Sci Law* 1999a;39:287-292.
- Ermenc B. Minimizing mistakes in clinical diagnosis. *J Forensic Sci* 1999b;44(4):810-13.
- Ermenc B. Comparison of the clinical and post mortem diagnoses of the causes of death. *Forensic Sci Int* 2000;114:117-119.
- Feinstein AR. Epidemiological and clinical challenges in reviving the necropsy. *Arch Pathol Lab Med* 1996;120:749-752.
- Fernandez-Segoviano P, Esteban A. The autopsy as a tool to monitor diagnostic error. *Intensive Care Med* 1999;25:343-344.

- Fernandez-Segoviano P, Lázaro A, Esteban A, Rubio JM, Iruretagoyena JR. Autopsy as quality assurance in the intensive care unit. *Crit Care Med* 1988;16(7):683-685.
- Gall EA. The necropsy as a tool in medical progress. *Bull N Y Acad Med* 1968;44:808-829.
- Gambino SR. The autopsy – the ultimate audit. *Arch Pathol Lab Med* 1984;108:444-445.
- Geller SA. Religious attitudes and the autopsy. *Arch Pathol Lab Med* 1984;108(6):494-496.
- Goldman L. Diagnostic advances vs the value of autopsy. *Arch Pathol Lab Med* 1984;108(6):501-505.
- Goldman L, Sayson R, Robbins S, Cohn LH, Bettman M, Weisberg M. The value of autopsy in three medical eras. *N Engl J Med* 1983;308(17):1000-1005.
- Goldstein B, Metlay L, Cox C, Rubenstein JS. Association of pre mortem diagnosis and autopsy findings in pediatric intensive care unit versus emergency department versus ward patients. *Crit Care Med* 1996;24:683-686.
- Gordon BM. Medical errors: creating the tension for change *Am J Health-Syst Pharm* 2001;58:908-912.
- Graber ML, Franklin N, Gordon R. Diagnostic error in internal medicine. *Arch Intern Med* 2005;165:1493-1499.
- Grade MH. Estudo do diagnóstico da causa básica de óbito, antes e após a necropsia, no HCFMRPUSP no período de 1990 a 1995, e análise comparativa com o final da década de 70 [dissertação]. Ribeirão Preto: Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo, 2003.
- Grade MH, Zucoloto S, Kajiwarra JK, Fernandes MT, Couto LG, Garcia SB. Trends of accuracy of clinical diagnoses of the basic cause of death in a university hospital. *J Clin Pathol* 2004;57:369–73.
- Gut AL, Ferreira ALA, Montenegro MR. Autopsy: quality assurance in the ICU. *Intensive Care Med* 1999;25:360-363.

Haber SL. Whither the Autopsy? Arch Pathol Lab Med 1996;120:714-717.

Haque AK, Patterson RC, Grafe MR. High autopsy rates at a university medical center - What has gone right? Arch Pathol Lab Med 1996;120:727-732.

Hasuo Y, Ueda K, Kiyohara Y, Wada J, Kawano H, Kato I, et al. Accuracy of diagnosis on death certificates for underlying causes of death in a long-term autopsy-based population study in Hisayama, Japan; with special reference to cardiovascular diseases. J Clin Epidemiol 1989;42(6):577-584.

Helmreich RL. On error management: lessons from aviation. BMJ 2000;320:781-785.

Hill RB. College of American Pathologists Conference XXIX on Restructuring Autopsy Practice for Health Care Reform: Summary. Arch Pathol Lab Med 1996;120:778-785.

Hill RB, Anderson RE. The autopsy – medical practice and public policy. Boston : Butterworth Publishers, 1988;294 p.

Hill RB, Anderson RE. The recent history of the autopsy. Arch Pathol Lab Med 1996;120:702-712.

Hutchins GM. College of American Pathologists Conference XXIX on Restructuring Autopsy Practice for Health Care Reform: Introduction. Arch Pathol Lab Med 1996 a;120:733.

Hutchins GM. Whither the autopsy?...To regional autopsy centers. Arch Pathol Lab Med 1996b;120:718.

Institute of Medicine. To err is human: building a safety health system. Washington, DC: National Academy Press, 1999.

Juvin P, Teissiere F, Brion F, Desmonts JM, Durlin M. Postoperative death and malpractice suits: is autopsy useful? Anesth Analg 2000;91(2):344-346.

Kaafarani HM, Itani KM. The surgeon's ego and reporting of clinical failure. Br J Surg 2005;92:119-122.

- Kajiwara JK. Comparação entre os diagnósticos de causas básicas de morte antes e após a autópsia no HCFMRPUSP entre 1978 e 1980[dissertação]. Ribeirão Preto: Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo, 1991.
- Kajiwara JK, Zucoloto S. A Autópsia na História. *Medicina, Ribeirão Preto* 1991;24(2):55-69.
- Kajiwara JK, Zucoloto S, Manço ARX, Muccillo G, Barbieri MA. Accuracy of clinical diagnoses in a teaching hospital: a review of 997 autopsies. *J Intern Med* 1993;234:181-187.
- Karsh BT. Beyond usability: designing effective technology implementation systems to promote patient safety. *Qual Saf Health Care* 2004;13:388-394.
- Kassirer JP. Teaching problem-solving-how are we doing? *N Engl J Med* 1995;332:1507-1509.
- Kaufman SR. Autopsy - A crucial component of human clinical investigation. *Arch Pathol Lab Med* 1996;120:767-770.
- King DW. Potential of the autopsy. *Arch Pathol Lab Med* 1984;108:439-443.
- King LS, Meehan MC. A history of autopsy. *Am J Pathol* 1973;73:514-544.
- Kirch W, Schaffii C. Misdiagnosis at a university hospital in 4 medical eras. *Medicine (Baltimore)* 1996;75(1):29-40.
- Kircher T, Nelson J, Burdo H. The autopsy as a measure of accuracy of the death certificate. *N Engl J Med* 1985;313:1263-1269.
- Kumar P, Angst DB, Taxy J, Mangurten HH. Autopsies in children: are they still useful? *Arch Pediatr Adolesc Med* 1998;152(6):558-563.
- Kumar P, Angst DB, Taxy J, Mangurten HH. Neonatal autopsies: a 10-year experience. *Arch Pediatr Adolesc Med* 2000;154(1):38-42.

- Kuperman GJ, Gibson RF. Computer physician order entry: Benefits, costs, and issues. *Ann Intern Med* 2003;139:31–39.
- Landfeld CS, Chren MM, Myers A, Geller R, Robbins S, Goldman L. Diagnostic yield of the autopsy in a university hospital and a community hospital. *N Engl J Med* 1988;318:1249-1254.
- Lardenoye JW, Kappetein AP, Peeters MP, Spaander PJ, Breslau PJ. Value of keeping records of mortality. *Eur J Surg* 2002;168:436–440.
- Lau DT, Kasper JD, Potter DEB, Lyles A, Bennett RG. Hospitalization and death associated with potentially inappropriate medication prescriptions among elderly nursing home residents. *Arch Intern Med* 2005;165:68-74.
- Leape LL. Error in Medicine. *JAMA* 1994;272:1851-1857.
- Leape LL, Woods DD, Hatlie MJ, Kizer KW, Schroeder SA, Lundberg GD. Promoting patient safety by preventing medical error. *JAMA* 1998;280(16):1444-1447.
- Lester H, Tritter JQ. Medical error: a discussion of the medical construction of error and suggestions for reforms of medical education to decrease error. *Med Educ* 2001;35:855-861.
- Lundberg GD. Medicine without the autopsy. *Arch Pathol Lab Med* 1984;108(6):449-454.
- Lundberg GD. College of American Pathologists Conference XXIX on restructuring autopsy practice for health care reform: let's make this autopsy conference matter. *Arch Pathol Lab Med* 1996;120:736-737.
- Margo CE. A pilot study in ophthalmology of inter-rater reliability in classifying diagnostic errors: an under investigated area of medical error. *Qual Saf Health Care* 2003;12:416-420.
- McNutt RA, Abrams R, Aron DC. Patient safety efforts should focus on medical errors. *JAMA* 2002;287(15):1997-2001.

- McPhee SJ. Maximizing the benefits of autopsy for clinicians and families - What needs to be done. *Arch Pathol Lab Med* 1996a;120:743-748.
- McPhee SJ. The autopsy: an antidote to misdiagnosis. *Medicine* 1996b;76:41-43.
- Meyer G, Lewin DI, Eisenberg J. To err is preventable: medical errors and academic medicine. *Am J Med* 2001;110(7):597-603.
- Mort TC, Yeston NS. The relationship of pre mortem diagnoses and post mortem findings in a surgical intensive care unit. *Crit Care Med* 1999;27:299-303.
- Nadrous HF, Afessa B, Pfeifer EA, Peters SG. The role of autopsy in the intensive care unit. *Mayo Clin Proc* 2003;78:947-950.
- Newton D, Coffin CM, Clark EB, Lowichik A. How the pediatric autopsy yields valuable information in a vertically integrated health care system. *Arch Pathol Lab Med* 2004;128:1239-1246.
- O'leary DS. Relating autopsy requirements to the contemporary accreditation process. *Arch Pathol Lab Med* 1996;120:763-766.
- O'neil AC, Petersen LA, Cook F, Bates DW, Lee TH, Brennan TA. Physician reporting compared with medical-record review to identify adverse medical events. *Ann Intern Med* 1993;119:370-376.
- Ong AW, Cohn SM, Cohn KA, Jaramillo DH, Parbhu R, McKenney MG, et al. Unexpected findings in trauma patients dying in the intensive care unit: results of 153 consecutive autopsies. *J Am Coll Surg* 2002;194(4):401-406.
- Organização Mundial de Saúde (OMS). Classificação estatística internacional de doenças e problemas relacionados à saúde (CID) - 10ª Revisão, 1989.
- Postman N. *Technopoly: The surrender of culture to technology*. New York: Vintage Books; 1992.

- Pellegrino ED. The autopsy - some ethical reflections on the obligation of pathologists, hospitals, families, and society. *Arch Pathol Lab Med* 1996;120:739-742.
- Reason J. Human error: models and management. *BMJ* 2000;320:768-770.
- Reiser SJ. *Medicine and the reign of technology*. Cambridge: Cambridge University Press; 1978.
- Ridley SA, Booth SA, Thompson CM and the Intensive Care Society's Working Group on Adverse Incidents. Prescription errors in UK critical care units. *Anaesthesia* 2004;59:1193-1200.
- Roulson J, Benbow EW, Hasleton PC. Discrepancies between clinical and autopsy diagnosis and the value of post mortem histology; a meta-analysis and review. *Histopathology* 2005;47:551-559.
- Salib E, Tadros G, Ambrose A. Autopsy in elderly psychiatric inpatients: a retrospective review of autopsy findings of deceased elderly psychiatric inpatients in North Cheshire 1980-1996. *Med Sci Law* 2000;40:20-27.
- Sanders RL. Medical technology: a critical perspective. *The Internet Journal of Medical Technology* 2004; Volume 2 Number 1. Disponível em: <http://www.ispub.com/ostia/index.php?xmlFilePath=journals/ijmt/vol2n1/medical.xml>.
- Saracci R. Is necropsy a valid monitor of clinical diagnosis performance? *BMJ* 1991;303:898-900.
- Seckinger DL. Our two-day exploration of the autopsy. *Arch Pathol Lab Med* 1996;120:719.
- Seow E, Lau G. Who dies at A & E? The role of forensic pathology in the audit of mortality in an emergency medicine department. *Forensic Sci Int* 1996;82:201-210.
- Setlow VP. The need for a national autopsy policy. *Arch Pathol Lab Med* 1996;120:773-777.
- Silverberg SG. The autopsy and cancer. *Arch Pathol Lab Med* 1984;108(6):476.

- Sinard JH, Blood DJ. Quality Improvement on an academic autopsy service. *Arch Pathol Lab Med* 2001;125:237-245.
- Shojania KG, Burton EC, McDonald KM, Goldman L. Changes in rates of autopsy-detected diagnostic errors over time: a systematic review. *JAMA* 2003;289:2849-2856.
- Shojania KG, Burton EC. The persistent value of the autopsy. *Am Fam Physician* 2004;69(11):2540-2542.
- Sonderegger-Iseli K, Burger S, Muntwyler J, Salomon F. Diagnostic errors in three medical eras: a necropsy study. *Lancet* 2000;335:2027-2031.
- Spencer FC. Human error in hospitals and industrial accidents: current concepts. *J Am Coll Surg* 2000;191(4):410-418.
- Spiliopoulou C, Papadodima S, Kotakidis N, Koutselinis A. Clinical diagnoses and autopsy findings: a retrospective analysis of 252 cases in Greece. *Arch Pathol Lab Med* 2005;129(2):210-214.
- Tai DYH, El-Bilbeisi H, Tewari S, Mascha EJ, Wiedemann HP, Arroglia AC. A study of consecutive autopsies in a medical ICU. *Chest* 2001;119:530-536.
- Temple WJ. Medical errors: a challenge to all physicians. *J Surg Oncol* 2004;88(3):111.
- Thornton CM, O'hara MD. A regional audit of perinatal and infant autopsies in Northern Ireland. *Br J Obstet Gynaecol* 1998;105:18-23.
- Trelstead RL, Amenta PS, Foran DJ, Smilow PC. The role for regional autopsy centers in the evaluation of covered deaths - survey of opinions of US and Canadian chairs of pathology and major health insurers in the United States. *Arch Pathol Lab Med* 1996;120:753-758.
- Tse GM, Lee JC. A 12-month review of autopsies performed at a university-affiliated hospital in Hong-Kong. *Hong Kong Med J* 2000;6(2):190-194.

- Tsujimura T, Yamada Y, Kubo M, Fushimi H, Kameyama M. Why couldn't an accurate diagnosis be made? An analysis of 1044 consecutive autopsy cases. *Pathol. Int.* 1999;49:408-410.
- Twigg SJ, McCrirrick A, Sanderson PM. A comparison of post mortem findings with post hoc estimated clinical diagnoses of patients who die in a United Kingdom intensive care unit. *Intensive Care Med* 2001;27:706-710.
- Vadillo M, Pujol RP, Corbella X, Gorriz T, Rabasa P, Bernat R. Improvements in clinical diagnostic accuracy after a 5-year systematic analysis of clinical and autopsy discrepancies. *Arch Pathol Lab Med* 2006;130:1261-1262.
- Valdez-Martínez E, Arroyo-Lunagómez E, Landero-López L. Concordância entre el diagnóstico clínico y el patológico por necropsias. *Salud Publica Mex* 1998;40:32-37.
- Voytovich AE, Rippey RM, Suffredini A. Premature conclusions in diagnostic reasoning. *J Med Educ* 1985;60:302-307.
- Xavier AC, Siqueira SA, Costa LJ, Mauad T, Nascimento-Saldiva PH. Missed diagnosis in hematological patients-an autopsy study. *Virchows Arch* 2005;446(3):225-231.
- Weingart SN, Wilson RM, Gibberd RW, Harisson B. Epidemiology of medical error. *BMJ* 2000;320:774-777.
- Wieman TJ, Wieman EA. A systems approach to error prevention in medicine. *J Surg Oncol* 2004;88:115-121.
- Wissler RW. The value of the autopsy for understanding cardiovascular disease. *Arch Pathol Lab Med* 1984;108(6):479.
- Wolf SH, Kuzel AJ, Dovey SM, Phillips RL. A string of mistakes: the importance of cascade analysis in describing, counting, and preventing medical errors. *Ann Fam Med* 2004;2:317-326.
- Zaitoun AM, Fernandez C. The value of histological examination in the audit of hospital autopsies: a quantitative approach. *Pathology* 1998;30(2):100-104.

Zarbo RJ, Baker PB, Howanitz PJ. The autopsy as a performance measurement tool - diagnostic discrepancies and unresolved clinical questions. *Arch Pathol Lab Med* 1999;123:191-198.

Zher KJ, Liddicoat JR, Salazar JD, Gillinov AM, Hruban RH, Hutchins GM, et al. The autopsy: still important in cardiac surgery. *Ann Thorac Surg* 1997;64:380-383.

Anexo de Publicação

Despite the modern medical technology the initial diagnosis is still the main source of error along the patient trajectory in a tertiary referral center hospital.

Marcio Henrique Carvalho Grade, Sérgio Zucoloto, Sérgio Britto Garcia *

SUMMARY

Background - Despite much progress in medical technology, assessment of its impact on diagnostic accuracy has remained a problematic issue. The role of autopsy in auditing clinical diagnosis performance is well demonstrated. We studied the quality assessment of medical diagnostics, based on detailed, ongoing comparison clinical diagnoses with autopsy findings.

Methods - Three hundred patients who had died at a tertiary-care teaching hospital in Brazil—100 in each from 1980, 1990, and 2000 were randomly selected. Diagnostic errors were analyzed retrospectively in order to verify the discrepancies between the clinical initial and final diagnosis and necropsy findings. The diagnostics were classified in major and minor discrepancies and concordance.

Findings - There was an increase in agreement between the final clinical diagnosis and the autopsy findings during the period studied (41 to 55%), due to an increase in concordances

of minor diagnostics. Nevertheless, there was no reduction of the major discrepancies that have maintained to a 21% rate. Furthermore, a high and constant concordance rate (80%) was verified between the initial and the final clinical diagnostic, despite a remarkable increase in the number of complementary exams (from 236 to 366). Interestingly, an even higher percentage of concordances occurred in the cases with major diagnostic discrepancies (89%).

Interpretation – There is a clear tendency for doctors not to change their initial clinical diagnostic, despite of the increasing use of modern diagnostic technology. The premature closing of the diagnosis seems to be the main factor related to the maintenance of these major discrepancies between the clinical diagnostic and necropsy findings.

Introduction

There is a very basic question in medical practice that has not been appropriately addressed: What is the impact of modern medical technology for diagnostic accuracy?

Most people are familiar with media reports on the benefits of medical technologies every day. However, it is not often that physicians have the opportunity to discuss what has been given up or lost as a result of using these technologies (1). Indeed, although much progress has been made in medical technology, assessment of its impact in diagnostic accuracy has remained a problematic issue. Interestingly, there has been no systematic comprehensive study of the sensitivity, specificity, cost-benefit ratio, and reliability of physical diagnosis relative to technologic diagnostic tools. This is an important issue because modernization in medical technology has led people to put high demands on the performance of medical practice in the last decades.

In fact, there has been little improvement on the overall rate of discrepancies between the clinical and autopsy diagnostics since the 1960s, despite the great advances in diagnostic techniques (2,3). This disappointing performance of the modern medical technological exams in improving the quality of the diagnostics has not been deeply evaluated and may have profound repercussions in the management of malpractice, raising ethical and economical discussions. Therefore instruments that access and measure the effectiveness and limitations of the technological support in the quality of the medical diagnosis are needed.

The role of autopsy in auditing clinical diagnosis performance has been shown by many investigators assessing the discrepancy between clinical and autopsy diagnostics and it is generally accepted that autopsy may be used as a tool for quality management to analyze

diagnostic discrepancies (4,5,6). Despite the fact that the autopsy service itself should be subjected to quality control to assure the reliability of the autopsy findings, including the quality of diagnoses made postmortem, the autopsy reports are usually considered as the golden standard of the diagnostic procedure (7). A systematic literature search for English-Language articles has shown a quite high median diagnostic error rate, that was of 23.5% for major errors (8) and studies are needed to assess how and when diagnostic errors occur so that this might improve diagnostic performance (9).

The clinical diagnostic has two major classical steps which are: 1 - the initial diagnostic, performed after anamnesis and physical exams, and 2 - the final diagnostic performed after the less invasive complementary exams (image exams, serological tests) and the more invasive complementary exams (like biopsies, exploratory surgeries). Errors may occur in both of these steps, but there have been very few systematic studies of the sensitivity and reliability of these steps of the clinical diagnostic and none of them have used autopsy as a golden standard. Therefore, we here propose a study to verify the quality assessment of medical diagnostics, based on detailed, ongoing comparison of three steps of the clinical diagnoses with autopsy diagnoses.

Methods

Providing medical care at the primary, secondary, tertiary *and* emergency level, the public Hospital of the Faculty of Medicine of Ribeirão Preto, University of São Paulo (HCFMRP), is a teaching and regional reference center. We had previously verified that the discrepancy rate between clinic and autopsy diagnostic in the HCFMRP is around 21%, which

is situated below the international average (8). In the HCFMRP the percentage of deceased patients submitted to autopsy is quite high. Usually more than 70% of the deceased patients are submitted to autopsy (10). The autopsy is performed by the pathologist trainee under direct supervision of the teaching staff of the Department of Pathology.

We retrospectively analyzed the medical and necropsy records of three hundred randomly selected patients admitted to the HCFMRP, divided into one hundred patients in 1980, 1990, and 2000 respectively. This sampling was chosen in accordance to previous studies (3,11). Only patients who died at the Hospital and underwent necropsy were considered for the study. The first time period (1980) was chosen because most of the modern image exams were available from that time on at the HCFMRP. A detailed and comprehensive review of the autopsy and clinical reports of each patient was performed. The study was approved by the Research Ethics Committee of the institution.

There were no restrictions to the selection of the patients in the study. The evaluation of the records included the discrimination of each patient's diagnostics in three different diagnostic steps: (a) the initial diagnostic (the diagnostic performed before the request of technological exams), (b) the final diagnostic (performed by the physician, after the death of the patient taking into account the results of all the complementary exams that led to specific treatments or conducts), and (c) the autopsy diagnostic. In sequence, correlations among the diagnostic steps were performed in accordance to the classification of Goldman and colleagues (11), modified by Battle and colleagues (12) and by Sonderegger-Iseli and colleagues (3), summarized at the Panel 1. The major diagnostic discrepancies (class I and II) had been considered when related to the basic cause of death or when they had directly contributed to death. The minor discrepancies were those that had not been related to the basic or direct

causes of death. The classification of the discrepancies was carried through by two medical researchers, an experienced general practitioner and a pathologist.

The diagnostic classified as class I, had been revised for a second time, considering the eventual necessity of a reclassification. The accuracy, sensitivity and specificity for the biggest clinical categories have been calculated. Accuracy was calculated as true positives or true negatives divided into all the cases. Sensitivity was defined as the ratio of true positives divided by true positives plus the false negatives. Specificity was defined as the ratio of true negatives divided by the addition of true negatives and false positives.

The image examinations had been discriminated in: radiological contrasted and non-contrasted, computerized tomography, endoscopy, biopsies, cintilography, ultrasonography (including echocardiography), cat scans and magnetic resonances. The results of discrepancies had been analyzed with the accurate test of Cochran-Armitage for the comparison of years ($p < 0,05$), as previously described (3).

All the cases that were considered as discrepancies class I were reevaluated and classified according to the type of flaws in exams that were performed in these patients, as in the following classes:

- 1- Exams that would have been essential for the correct diagnostic were not requested.
- 2- Appropriate exams were performed but the interpretation of their results was flawed.
- 3- Appropriate exams were performed but their results were false negatives due to insufficient sensitivity or false positives due to a lack of specificity.
- 4 –Not classifiable cases with the available data.

Finally, the data had been expressed in medium, average and ratio.

Results

In our sampling, there has been an increase in the age and in the number of patients with more than two previous internments from 1980 to 2000 (Table 1). The main percentage of deaths was due to the cardiovascular diseases, with a decrease from 43% of the cases in 1980 to 30% in 2000. The second cause of deaths was the neoplasias, which increased from 12% in 1980 to 23% in 2000. The prevalence of other relevant causes of death, such as illnesses of the digestive, respiratory, endocrine/metabolic systems and infectious diseases did not present significant changes in the period of the study. The microscopy elucidated, completed or modified the major diagnoses in 7% of the cases in 1980, 4,9% in 1990 and 7,3% in 2000.

There was a remarkable increase in the number of complementary exams such as endoscopy, biopsies, ultrasonography, tomography and cat scans in the period of the study, with a reduction of the number of radiological examinations and exploratory surgeries (Table 2). As shown in table 3, there was a discrete improvement of the specificity and accuracy for the exams related to the cardiovascular diseases, increase in the diagnostic sensitivity of neoplasias and reductions in the sensitivity for the infectious diseases and illnesses of the digestive and respiratory systems.

Figure 1 presents the correlations among the diagnostics: Correlation A (between final clinical diagnostic and the autopsy findings), Correlation B (between the initial clinical diagnostic and the autopsy) and Correlation C (between initial and final clinical diagnostic). In figures 1A and 1B a gradual increase in the percentage of agreement for the correlations A and B can be noticed ($p=0.05$ and $p=0.03$, respectively), with a concomitant reduction of the minor discrepancies. However, the major discrepancies remained unchanged in the three periods.

There were no significant changes in the concordances and discrepancies of Correlation C (Fig.1C) along the time periods of the study, with an exception of an increase in the Class I discrepancies in the last analyzed period. The agreement between the initial and final clinical diagnosis has remained around 80% in the three evaluated periods. A further evaluation has shown and even higher concordance rate (89%) between the initial and final clinical diagnostic (Correlation type C) for the patients with Class I discrepancy between the final clinical and autopsy diagnostic (Correlation type A).

Tables 4 and 5 respectively detailed the cases with discrepancies of Class I and II. One notices that the gastrointestinal illnesses are the main group with Class I discrepancies, and unrecognized gastric peptic ulcer is the main single major diagnostic flaw. The respiratory diseases are responsible for the biggest Class II discrepancies, with pneumonias appearing to be the main unrecognized diagnostic. The cardiovascular diseases are the second cause of Class I and II discrepancies.

As shown in table 6, in the cases with major diagnostic discrepancies the most frequent diagnostic flaws are of the Class 1, followed by the Class 2. No case of the Class 3 flaw has been observed.

Discussion

In our study cardiovascular diseases were the most common, followed by the neoplasias and the gastrointestinal diseases. There was a discrete increase in the diagnostic accuracy and sensitivity of cardiovascular diseases and a low sensitivity for the digestive diseases, mainly due to the misdiagnosis of perforated peptic ulcer. Digestive bleeding, was the

most common cause of class I discrepancies, similarly to the findings of Tse and Lee (13). It may be considered that this type of diagnosis can be difficult to carry through in patients with difficulty in communication and must be taken into account for the patients submitted to severe stress. Bronchopneumonia was other frequent diagnosis between the discrepancies, this is expected because it is one of the most common causes of hospital internments and when not related with the internment diagnostic it usually is responsible for the aggravation of the general clinical status of the patients.

Our study presented a significant increase in the diagnostic agreements but this was only due to a reduction of the minor discrepancies. This reduction may be related to an increase in the number of diagnostic procedures and in the percentage of patients with more than 2 previous internments.

In the comparison between the final clinical diagnostic and autopsy (Correlation type A), the major discrepancies presented a remarkable stability over time during the three decades. This result corroborates previous studies, as much for the general analyses as for the distinguished patients, the restricted groups of diseases and environments (14). This is in accordance to Goldmand et al. (15) who proposed that an over-reliance on the new diagnostic procedures may have occasionally contributed directly to the flaws in major diagnoses.

Interestingly, the agreement between the initial and final clinical diagnosis (Correlation type C) remained around 80% in the three periods. This strongly suggests that the doctors show an important resistance of changing their initial diagnostic hypothesis. This suggestion was reinforced by the further evaluation that, in the cases of class I discrepancies, there was an even higher agreement between the initial and final clinical diagnosis (89%). Moreover, we did not find differences between the number of complementary exams between the groups of

patients with diagnostic agreement and major discrepancies.

Sonderegger-Iseli et al. (3) have proposed a framework for errors that may be useful to explain our findings. These authors suggest that the premature closing of the diagnostic was one of main reasons responsible for the diagnostic discrepancy, been related to the generation of the hypothesis and request of exams, inadequate interpretation of the tests or the lack of suspicion for other possible diagnostics (16). Graber et al (9) have studied one hundred diagnostic errors involving internists, which were identified through autopsy discrepancies. Premature closure, i.e., the failure to continue considering reasonable alternatives after an initial diagnosis was reached, was found to be the most common cause of diagnostic error.

Therefore, the improvement of the quality of the initial diagnostics is crucial for the improvement of the overall diagnostic accuracy. Vadillo et al. (17) have obtained a reduction of the major discrepancies after the systematic analysis of pre-autopsy diagnoses stated by the clinicians involved with the patients and the resulting analysis of clinical and autopsy discrepancies. This continuous feedback between clinicians and pathologists with systematic revision demonstrated a capability of diminishing the diagnostic discrepancies.

Our findings point to the need of a more comprehensive understanding of the role and limits of modern technology in the diagnostic accuracy. Over the past hundred years one can observe that there has been an increased faith in science and technology, and on the reliance in scientific quantitative analysis conducted in the laboratory over the more qualitative diagnosis by the physician (18). This is in accordance with the view that throughout the 20th century, a change occurred in which our reliance on technology elevated it from a means to an end, whereby technology no longer emerged out of a need to solve a problem, but rather as a manifest destiny based on the belief that if it could be invented, it should (19). This change, in

our point of view of technology, subsequently discouraged 20th century physicians from continuing to develop the art of diagnosis, now considered imprecise and time-consuming (19). Sanders (1) has appropriately put medical technology in perspective when he proposed an answer to the question whether technology lead to better medicine. He proposed that the answer is “a simple; *Yes* and *No*. The answer is *Yes* if one can look at all the wonderful and amazing things that can be done with technology (diagnoses, procedures, minimized pain, outpatient surgeries, etc.). On the other hand, the answer is *No* if one considers the many needless procedures done because they can be done or those procedures that can potentially lead to other problems (i.e. X-rays and other iatrogenics, treatment induced illnesses)”. Sanders considers that “technology not only distances the physician from the patient, but also distances the physician from himself, his personal history, and the history of his profession”, and cited Dr. Robert J. Weiss of Columbia University as supporting this view by stating, “For a profession which considers itself scientific, we have approached the issue of the appropriate use of technology and treatment in a most unscientific fashion... Not only have we not been scientific, we have not been humanitarian”

Our data may also be used to approach the difficult subject of a “necessary fallibility” of physicians (20). Even accepting that there should be a continuous search for an “error-free medicine”, it is necessary to discuss the problem of the natural fallibility of any process, where the diagnostic process with its innumerable variables is also subject to fallibility. Irrational faith in technology may lead us to forget these aspects and this may have serious consequences in clinical, economical, ethical and forensic paradigms.

Conclusions

Our study presented a significant increase in the agreements between the final clinical diagnostic and autopsy findings, but this was only due to a reduction of the minor discrepancies. This reduction may be related to an increase in the number of diagnostic procedures and in the percentage of patients with more than 2 previous interments. The major discrepancies had presented a remarkable stability over time during three decades. The agreement between the initial and final clinical diagnosis has remained around 80% in the three periods. This strongly suggests that the doctors show an important resistance of changing their initial diagnostic hypothesis, which is reinforced by the further evaluation that, in cases with class I discrepancies, there was an even higher agreement between the initial and final clinical diagnosis. Moreover, we did not find differences between the number of complementary exams between the groups of patients with diagnostic agreement and major discrepancies. Thus, the increase in the number of more modern examinations had basically an impact for the reduction of the minor discrepancies, without affecting the major ones. Our study suggests that the premature closing of the diagnosis is greatly responsible for the diagnostic errors.

REFERENCES

- 1) Sanders RL. Medical Technology: A Critical Perspective. The Internet Journal of Medical Technology 2004; Volume 2 Number 1. Available from: <http://www.ispub.com/ostia/index.php?xmlFilePath=journals/ijmt/vol2n1/medical.xml>.
- 2) Editorial. When primum non nocere fails. Lancet 2000; 355: 2007.
- 3) Sonderegger-Iseli K, Burger S, Muntwyler J, Salomon F. Diagnostic errors in three medical eras: a necropsy study. Lancet 2000; 335: 2027–31.
- 4) Newton D, Coffin CM, Clark EB, Lowichik A. How the Pediatric Autopsy Yields Valuable Information in a Vertically Integrated Health Care System. Arch Pathol Lab Med. 2004; 128: 1239–1246.
- 5) Lardenoye JW, Kappetein AP, Peeters MP, Spaander PJ, Breslau PJ. Value of Keeping Records of Mortality. Eur J Surg. 2002; 168: 436–440.
- 6) Nadrous HF, Afessa B, Pfeifer EA, Peters SG. The role of autopsy in the intensive care unit. Mayo Clin Proc. 2003; 78: 947-50.
- 7) Bayer-Garner IB, Fink LM, Lamps LW. Pathologists in a teaching institution assess the value of the autopsy. Arch Pathol Lab Med. 2002; 126: 442–7.
- 8) Shojania KG, Burton EC, McDonald KM, Goldman L. Changes in rates of autopsy-detected diagnostic errors over time: a systematic review. JAMA. 2003; 289: 2849-56.
- 9) Graber ML, Franklin N, Gordon R. Diagnostic error in internal medicine. Arch Intern Med. 2005; 165: 1493-9.

- 10) Grade MH, Zucoloto S, Kajiwara JK, Fernandes MT, Couto LG, Garcia SB. Trends of accuracy of clinical diagnoses of the basic cause of death in a university hospital. *J Clin Pathol.* 2004; 57: 369–73.
- 11) Goldman L, Sayson R, Robbins S, Cohn LH, Bettman M, Weisberg M. The value of autopsy in three medical eras. *N Engl J Med.* 1983; 308: 1000-5.
- 12) Battle M, Pathak D, Humble CG, et al. Factors influencing discrepancies between premortem and post-mortem diagnoses. *JAMA.* 1987; 258: 339–44.
- 13) Tse GM, Lee JC. A 12-month review of autopsies performed at a university affiliated hospital in Hong-Kong. *Hong Kong Med J.* 2000; 6: 190–4.
- 14) Roulson J, Benbow EW & Hasleton PC. Discrepancies between clinical and autopsy diagnosis and the value of post mortem histology; a meta-analysis and review. *Histopathology* 2005; 47: 551-9.
- 15) Goldman L. Diagnostic advances vs the value of autopsy. *Arch Pathol Lab Med.* 1984; 108: 501–5.
- 16) Voytovich AE, Rippey RM, Suffredini A. Premature Conclusions in Diagnostic Reasoning. *J Med Educ.* 1985; 60: 302-7.
- 17) Vadillo M, Pujol RP, Corbella X, Gorriz T, Rabasa P, Bernat R. Improvements in Clinical Diagnostic Accuracy After a 5-Year Systematic Analysis of Clinical and Autopsy Discrepancies. *Arch Pathol Lab Med.* 2006; 130: 1261–2.
- 18) Reiser SJ. *Medicine and the reign of technology.* Cambridge: Cambridge University Press; 1978.
- 19) Postman N. *Technopoly: The surrender of culture to technology.* New York: Vintage Books; 1992.
- 20) Anderson RE, Hill RS, Key CR. The sensitivity and specificity of clinical diagnostics during five decades. Toward an understanding of necessary fallibility. *JAMA.* 1989; 261: 1610-7.

Panel 1: Description of discrepancy classes according to Goldman and colleagues¹¹ modified by Battle and colleagues¹² and by Sonderegger-Iseli and colleagues³.

Major discrepancies

Class I

Discrepancies in major diagnoses where the knowledge of diagnosis before death would have led to changes in management that could have prolonged survival or cured the patient (eg, pulmonary infarction treated as pneumonia, fungal pneumonia treated as bacterial infection).

Class II

Discrepancies in major diagnoses which detection before death would not have changed survival even with the correct treatment (e.g. biventricular cardiac insufficiency due to severe aortic stenosis with missed pulmonary emboli, correctly treated bacterial sepsis with multi-organ failure because of the unrecognized postoperative cervical osteomyelitis in a patient with rheumatoid arthritis).

Minor discrepancies

Class III

Discrepancies in minor diagnoses not directly related to the cause of death, but with symptoms that should have been treated or would have eventually affected prognosis (e.g. pulmonary carcinoma in a patient with ruptured infra-renal aortic aneurysm).

Class IV

Discrepancies in minor diagnoses, occult or non-diagnosable, but with a possible epidemiological or genetic importance (e.g. symptomless gallstones, goitre).

Non discrepancy

Class V

Non-discrepant diagnoses.

Non-classifiable cases

Class VI

Patient died immediately after admission with no diagnostic procedures, or refused any diagnostic procedures or treatment. Necropsy was unsatisfactory with no clear findings and no diagnosis could be established after review of clinical and necropsy data.

	1980(n=100)	1990(n=100)	2000(n=100)
Male/Female	55/45	53/47	47/53
Age (median/media years)	51,5/47,4	50,5/47,1	65/56,7
Length of stay (median/media days)	8,5/16,1	6/15,3	7/13,5
Patients without previous admissions	64%	71%	60%
Patients with 1 previous admission	24%	12%	15%
Patients with ≥ 2 previous admissions	12%	17%	25%

Table 1 – General data of the studied patients

Procedure	1980	1990	2000	Total
Non-contrast radiological procedures	120	118	117	355
Contrast radiological procedures	39	25	14	78
Endoscopies	13	10	32	55
Biopsies	23	43	57	123
Surgical explorations	8	6	3	17
Scintigraphy	11	9	10	30
Ultrasound (including echocardiography)	2	35	65	102
Computed tomography	1	14	59	74
Magnetic resonance imaging	0	0	9	9
Total	217	260	366	843
Patients not submitted to exams	6	6	2	14

Table 2 – Number of diagnostic procedures in the years of the study

Diagnostic group	Accuracy			Sensitivity			Specificity		
	1980	1990	2000	1980	1990	2000	1980	1990	2000
Cardiovascular diseases	90%	91%	93%	78,5%	79%	76%	86%	90%	93%
Neoplasias	96%	96%	96%	75%	83%	83%	95%	95%	95%
Digestive tract diseases	98%	97%	96%	67%	67%	50%	98%	99%	98%
Respiratory diseases	99%	95%	95%	100%	100%	70%	99%	95%	96%
Infections	97%	99%	97%	70%	100%	80%	96%	99%	98%
Mental diseases	100%	96%	97%	100%	60%	67%	100%	99%	97%

Table 3 – Accuracy, sensitivity and specificity for diagnostic groups

System	Nº (%)	Disease
Digestive tract	15 (40)	Perforated ulcer (7), other causes of peritonitis (3), surgical iatrogeny (2), pancreatitis (1), cleft carcinoma (1), esophagus Paracoccidioidomycosis (1).
Circulatory	10 (27)	Myocardial infarction (2), cardiac insufficiency + acute pulmonary edema (2), Chagas disease (1), ischaemic disease (1), aortic dissection aneurysm (1), endocarditis (1), cardiac tamponade (1), cor pulmonale(1).
Respiratory	9 (24)	Pneumonia (4), pulmonary embolism (1), pneumothorax (1), foreign body (1), tuberculosis (1), pulmonary haemorrhage (1).
Nervous	2 (6)	Bacterial meningitis (1), tuberculosis meningitis (1)
Genitourinary	1 (3)	Acute renal insufficiency (1)
Total	37 (100)	

Table 4– Diagnoses of the class I discrepancies

System	N° (%)	Disease
Respiratory	12 (46)	Pneumonia (6), embolism (2), pulmonary carcinoma (1), pulmonary fibrosis (1), pulmonary haemorrhage (1), chronic obstructive pulmonary disease (1).
Circulatory	9 (35)	Cerebrovascular accident (4), cardiac ischaemic disease (2), aortic and mitral disease (1), Chagas disease (1), hypertrophic cardiomyopathy (1).
Digestive tract	5 (19)	Peritonitis (2), surgical iatrogeny (1), massive bowel necrosis (1), under-nutrition (1).
Total	26 (100)	

Table 5 –Diagnoses of the Class II discrepancies

Classes of flaws	1980	1990	2000	Total
1*	6 (54.5)	8 (61.5)	7 (53.8)	21 (56.8)
2†	0	2 (15.4)	4 (30.8)	6 (16.2)
3‡	0	0	0	0
4§	5 (45.5)	3 (23.1)	2 (15.4)	10 (27.0)
Total	11	13	13	37

Table 6 – Classes of flaws in exams that were performed in patients with Class I major diagnostic discrepancies in 1980, 1990 and 2000.

* - Exams that would have been essential for the correct diagnostic were not requested.

† - Appropriate exams were performed but the interpretation of their results was flawed.

‡ - Appropriate exams were performed but their results were false negatives due to insufficient sensitivity or false positives due to a lack of specificity.

§ - Not classifiable cases with the available data.

Figure 1 – Distribution of classes of discrepancies in a percentual comparison between final clinical diagnosis x autopsy (correlation A), initial clinical diagnosis x autopsy (correlation B), and initial x final clinical diagnosis (correlation C)