Chapter 5

Bayesian Analysis

Shuji Ogino and Robert B. Wilson

Introduction

The purpose of this chapter is to describe basic and general principles of Bayesian analysis for molecular pathologists. Thomas Bayes first described the theorem named after him in an essay on "the doctrine of chances," published posthumously in 1763, and republished in 1958.[1] Analyses based on Bayes' theorem are routinely applied to calculate probabilities in a wide variety of circumstances, not limited to medicine or genetics. In molecular pathology, Bayesian analysis is commonly used to calculate genetic risk, incorporating population data, pedigree information, and genetic testing results. First, Bayesian analysis will be introduced with two simple, concrete examples. In subsequent sections, the general principles illustrated by these examples are discussed and applied to more complex scenarios. For more in-depth treatments, the reader is referred to *Introduction to Risk Calculation in Genetic Counseling* by Young[2] and *The Calculation of Genetic Risks* by Bridge[3] as well as several articles on genetic risk assessment that include advanced Bayesian analyses, particularly for spinal muscular atrophy (SMA)[4,5] and cystic fibrosis (CF).[6–9]

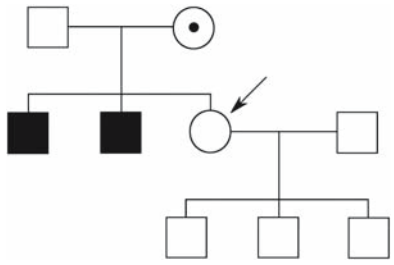
Bayesian Analysis Using Pedigree Information

In the pedigree shown in Figure 5-1a, the two brothers of the consultand (indicated by the arrow) have Kennedy disease (X-linked spinal and bulbar muscular atrophy; Online Mendelian Inheritance in Man [OMIM; database online] #300377), which is caused by a CAG trinucleotide expansion in the androgen receptor (*AR*) gene (OMIM #310200). Because both of the consultand's brothers are affected, we can assume that the consultand's mother is an obligate carrier. Before taking into account her three unaffected sons, the consultand's carrier risk is 1/2, since there is a 1/2 chance that she inherited the mutant X chromosome from her mother. If we take into account that the consultand has three unaffected sons, how does her carrier risk change?

Bayesian analysis starts with mutually exclusive hypotheses. In this example, there are two: that the consultand is a carrier, and that the consultand is a noncarrier. Setting up a table with separate columns for each hypothesis facilitates Bayesian analyses, as shown in Figure 5-1b for this case. The first row of the table comprises the prior probability for each hypothesis. In this example, the prior probabilities are the probability that the consultand is a carrier (1/2), and the probability that she is a noncarrier (also 1/2), *prior* to taking into account the subsequent information that she has three unaffected sons.

The second row of the table comprises the conditional probability for each hypothesis. The conditional probability for each hypothesis is the probability that the subsequent information would occur if we assume that each hypothesis is true. In this example, the subsequent information is that the consultand has three unaffected sons. Thus, the conditional probabilities are the probability that the consultand would have three unaffected sons under the assumption (or condition) that she is a carrier, and the probability that she would have three unaffected sons under the assumption (or condition) that she is a noncarrier. If we assume that she is a carrier, the probability that she would have three unaffected sons is $1/2 \times 1/2 \times 1/2 = 1/8$. This is because she would have to have passed the normal X chromosome three times in succession, each time with a probability of 1/2. If we assume that she is a noncarrier, the probability that she would have three unaffected sons approximates 1, since only in the event of a rare de novo mutation would a noncarrier have an affected son. Thus, the conditional probabilities in this example are 1/8 and 1 (Figure 5-1b).

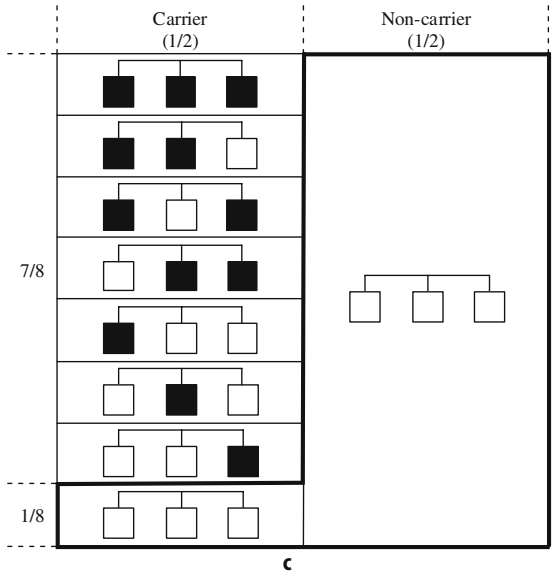
The third row of the table comprises the joint probability for each hypothesis, which is the product of the prior and conditional probabilities for each hypothesis. For the first hypothesis in this example, that the consultand is a carrier, the joint probability is the prior probability that she is a carrier, multiplied by the conditional probability that a



a

	Hypothesis	
	Carrier	Non-carrier
Prior Probability	1/2	1/2
Conditional Probability (of three normal sons)	1/8	~1
Joint Probability	1/16	1/2
Posterior Probability	(1/16) / (1/16 +1/2) = 1/9	(1/2) / (1/16 + 1/2) = 8/9

b



c

Figure 5-1. (a) Pedigree of a family with individuals affected with Kennedy disease (see text). Consultand is indicated by an arrow. (b) Bayesian analysis for the consultand in Figure 5-1a. (c) Schematic representation of the Bayesian analysis of Figure 5-1b. Pedigrees shown in the rectangles represent all possible disease status outcomes for the third generation of the pedigree in Figure 5-1a, given the carrier or noncarrier status of the consultand. Each small rectangle to the left represents one sixteenth of the total area. (See text for full description.)

carrier would have three normal sons, which in this case is $1/2 \times 1/8 = 1/16$ (Figure 5-1b). For the second hypothesis in this example, that the consultand is a noncarrier, the joint probability is the prior probability that she is a noncarrier, multiplied by the conditional probability that a noncarrier would have three normal sons, which in this case is $1/2 \times 1 = 1/2$ (Figure 5-1b).

The fourth row of the table comprises the posterior probability for each hypothesis. The posterior probability for each hypothesis is the probability that each hypothesis is true after (or posterior to) taking into account both prior and subsequent information. The posterior probability for each hypothesis is calculated by dividing the joint probability for that hypothesis by the sum of all the joint probabilities. In this example, the posterior probability that the consultand is a carrier is the joint probability for the first hypothesis (1/16), divided by the sum of the joint probabilities for both hypotheses ($1/16 + 1/2 = 9/16$), or $1/16 \div 9/16 = 1/9$. The posterior probability that the consultand is a noncarrier is the joint probability for the second hypothesis ($1/2 = 8/16$), divided by the sum of the joint probabilities for both hypotheses ($1/16 + 1/2 = 9/16$), or $8/16 \div 9/16 = 8/9$. Thus, taking into account the prior family history, and the subsequent information that the consultand has three unaffected sons, the probability that the consultand is a carrier is 1/9 (Figure 5-1b).

The preceding example is illustrated graphically in Figure 5-1c. The total area represents the total prior probabilities. The left half represents the prior probability that the consultand is a carrier (1/2), and the right half represents the prior probability that she is a noncarrier (also 1/2). Under the hypothesis that the consultand is a carrier, there are eight possibilities, comprising all the permutations of zero, one, two, or three affected sons. The area of the small rectangle that contains three unshaded squares (for three unaffected sons) comprises one eighth of the left half and represents the conditional probability of three normal sons under the hypothesis that the consultand is a carrier. The area of this small rectangle is one sixteenth of the total area and therefore also represents the joint probability that the consultand is a carrier (1/2), and that as a carrier she would have three normal sons (1/8), or $1/2 \times 1/8 = 1/16$.

Under the hypothesis that the consultand is a noncarrier, there is essentially only one possibility, which is that all three sons are unaffected. The area of the larger rectangle that contains the pedigree with three unshaded squares (for three unaffected sons) comprises all of the noncarrier half and represents the conditional probability of three normal sons under the hypothesis that the consultand is a noncarrier. The area of this larger rectangle is one half of the total area and therefore also represents the joint probability that the consultand is a noncarrier (1/2), and that as a noncarrier she would have three normal sons (~1), or $1/2 \times 1 = 1/2$. The reverse-L-shaped box, which is demarcated by a bold line, represents the sum of the joint probabilities, or nine sixteenths of the total area.

Because the consultand has three unaffected sons, the area of the reverse-L-shaped box represents the only component of the prior probabilities needed to determine the posterior probability that the consultand is a carrier. Taking into account that all three of the consultand's sons are unaffected, Bayesian analysis allows us to *exclude* 7/16 of the prior probabilities, those that include one or more affected sons, from consideration. (Note that this explains why the joint probabilities sum to less than 1.) The posterior probability that the consultand is a carrier is therefore the area of the small rectangle with three unshaded squares (for three unaffected sons) divided by the area of the entire reverse-L-shaped box, which represents the only probabil-

ities relevant to the consultand's risk, or 1/16 ÷ 9/16 = 1/9. Likewise, the posterior probability that the consultand is a noncarrier is the area of the larger rectangle with three unshaded squares (for three unaffected sons) divided by the area of the entire reversed-L-shaped box, or 8/16 ÷ 9/16 = 8/9.

Bayesian Analysis Using Genetic Test Results

In the second example, information from a test result modifies the prior risk. In the pedigree shown in Figure 5-2a, the consultand is pregnant with her first child and has a family history of CF (OMIM #219700). CF is caused by mutations in the cystic fibrosis transmembrane conductance regulator gene (*CFTR*; OMIM #602421). The consultand is an unaffected European Caucasian and her brother died years earlier of complications of CF. She undergoes carrier testing for the 23 mutations recommended by the American College of Medical Genetics (ACMG) CF screening guidelines,[10–12] which detects approximately 90% of disease alleles in European Caucasians. The consultand tests negative for all 23 mutations. What is her carrier risk after testing?

As in the first example, the two hypotheses are that the consultand is a carrier and that she is a noncarrier. The prior probability that she is a carrier is 2/3. Because the consultand is unaffected, she could not have inherited disease alleles from both parents. Thus, she either inherited a disease allele from her mother or father, or she inherited only normal alleles; in two of these three scenarios she would be a carrier (shown in Figure 5-2b). The prior probability that the consultand is a noncarrier is 1/3 (Figure 5-2c).

As in the first example, the conditional probability for each hypothesis is the probability that the subsequent information would occur if we assume that each hypothesis is true. In this example, the subsequent information is that the consultand tests negative for all 23 mutations. Thus, the conditional probabilities are the probability that the consultand would test negative under the assumption (or condition) that she is a carrier, and the probability that she would test negative under the assumption (or condition) that she is a noncarrier. If we assume that she is a carrier, the probability that she would test negative is 1/10, since the test detects 90% of European Caucasian disease alleles or carriers. If we assume that she is a noncarrier, the probability that she would test negative approximates 1. Thus, the conditional probabilities in this example are 1/10 and 1 for the carrier and noncarrier hypotheses, respectively (Figure 5-2c).

As in the first case, the joint probability for each hypothesis is the product of the prior and conditional probabilities for that hypothesis. For the first hypothesis in this example, that the consultand is a carrier, the joint probability is the prior probability that she is a carrier (2/3) multiplied by the conditional probability that a carrier of European Caucasian ancestry would test negative (1/10), or 2/3 × 1/10 = 1/15 (Figure 5-2c). For the second hypothesis in this example, that the consultand is a noncarrier, the

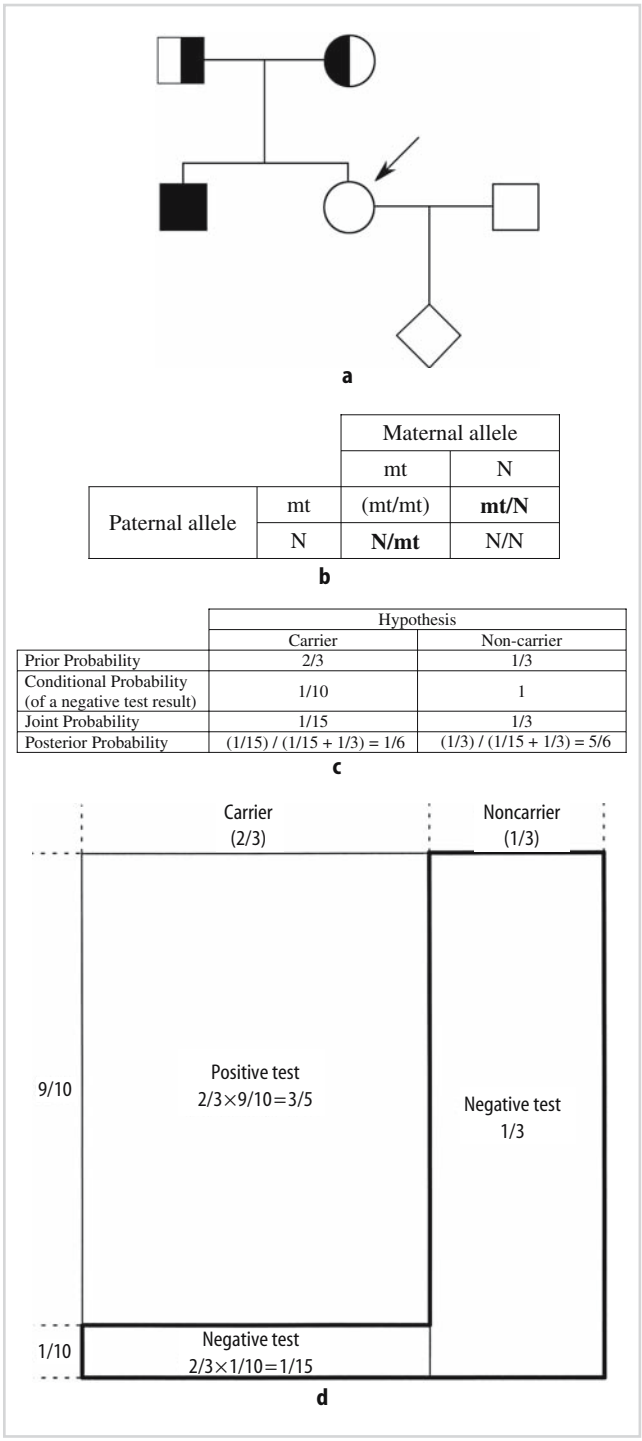


Figure 5-2. (a) Pedigree of a family with an individual affected with CF (see text). Consultand is indicated by an arrow. (b) Possible genotypes of the sibling (consultand in this case) of the affected child prior to genetic testing. The mt/mt genotype (in parentheses) is excluded based on the fact that the consultand is unaffected. mt, mutant; N, normal. (c) Bayesian analysis for the consultand in Figure 5-2a. (d) Schematic representation of the Bayesian analysis of Figure 5-2c (see text).

joint probability is the prior probability that she is a noncarrier (1/3) multiplied by the conditional probability that a noncarrier would test negative (1), or $1/3 \times 1 = 1/3$ (Figure 5-2c).

Finally, the posterior probability is calculated for each hypothesis by dividing the joint probability for that hypothesis by the sum of all the joint probabilities. In this example, the posterior probability that the consultand is a carrier and tests negative for 23 CF mutations is the joint probability for the first hypothesis (1/15) divided by the sum of the joint probabilities for both hypotheses ($1/15 + 1/3 = 2/5$), or $1/15 \div 2/5 = 1/6$ (Figure 5-2c). The posterior probability that the consultand is a noncarrier and tests negative for 23 CF mutations is the joint probability for the second hypothesis (1/3) divided by the sum of the joint probabilities for both hypotheses (2/5), or $1/3 \div 2/5 = 5/6$ (Figure 5-2c).

The preceding example is illustrated graphically in Figure 5-2d. The total area represents the total prior probabilities. The left two thirds represents the prior probability that the consultand is a carrier, and the right third represents the prior probability that the consultand is a noncarrier. Under the hypothesis that the consultand is a carrier, there are two possibilities for the test result: positive or negative. The area of the small rectangle on the lower left comprises one tenth of the 2/3 carrier region of the figure and represents the conditional probability of a normal test result under the hypothesis that the consultand is a carrier. The area of this small rectangle is $1/10 \times 2/3 = 1/15$ of the total probabilities area and therefore also represents the joint probability that the consultand is a carrier (2/3) and that as a European Caucasian carrier she would test negative for all 23 mutations (1/10), or $2/3 \times 1/10 = 1/15$ (Figure 5-2d).

Under the hypothesis that the consultand is a noncarrier, there is essentially only one possibility for the test result, which is negative. The area of the rectangle that comprises all of the 1/3 noncarrier region represents the conditional probability of a negative test result under the hypothesis that the consultand is a noncarrier. The area of this rectangle is one third of the total area and therefore also represents the joint probability that the consultand is a noncarrier (1/3), and that as a noncarrier she would test negative (~1), or $1/3 \times 1 = 1/3$. The reverse-L-shaped box, which is demarcated by a bold line, represents the sum of the joint probabilities, or 2/5 (= 1/3 + 1/15) of the total area.

Because the consultand tested negative, the area of the reverse-L-shaped box represents the only component of the prior probabilities needed to determine the posterior probability that the consultand is a carrier. Taking into account that she tested negative, Bayesian analysis allows us to *exclude* 3/5 of the prior probability, that portion comprising a positive test result, from consideration. (Note, again, that this explains why the joint probabilities sum to less than 1.) The posterior probability that the consultand is a carrier is therefore the area of the small rectangle at the lower left divided by the area of the reverse-L-shaped box, which represents the only probabilities relevant to the consultand's risk, or $1/15 \div 2/5 = 1/6$. Likewise, the posterior probability that the consultand is a noncarrier is the area of the larger rectangle on the right divided by the area of the reverse-L-shaped box, or $1/3 \div 2/5 = 5/6$.

Simple Bayesian Analyses Generalized: Carrier Versus Noncarrier

The preceding Bayesian analyses can be generalized as in Table 5-1. Note that if the correct prior and conditional probabilities can be determined, the rest is simple calculation. Setting up a spreadsheet, as in Table 5-1, facilitates clinical Bayesian analyses.

A very common application of Bayesian analysis in molecular pathology is to calculate carrier risk after a negative test result, as in the second example, above. The need to calculate carrier risk in this scenario stems from the fact that the sensitivity of most carrier tests is, at present, less than 100%; therefore, a negative test result decreases, but does not eliminate, carrier risk. Hypothesis 1 in this scenario is that the consultand is a carrier, and Hypothesis 2 is that the consultand is a noncarrier (Table 5-1). The prior carrier probability ("A" in Table 5-1) depends on whether there is a family history, and if there is, on the relationship of the consultand to the affected family member as shown by the family pedigree. In the absence of a family history, the prior carrier probability is the population carrier risk for that disease. In the case of CF and some other diseases, the appropriate population risk depends on the ethnicity of the consultand. The conditional probabilities ("C" and "D" in Table 5-1) are 1 minus the test sensitivity, and the test specificity, respectively. The remainder of the table is completed through calculation, with the posterior probabilities ("G" and "H" in Table 5-1) representing 1 minus the negative predictive value, and the negative predictive value, respectively. This is shown schematically in Figure 5-3.

For illustration, suppose in the second example above (Figure 5-2) that the consultand's husband is Ashkenazi Jewish, that he has no family history of CF, and that he tests negative for all 23 mutations in the ACMG screening guidelines panel. What is his carrier risk? The carrier risk in Ashkenazi Jewish populations, and therefore the husband's prior carrier risk in the absence of a family history, is approximately 1/25 ("A" in Table 5-1). Thus, his prior probability of being a noncarrier is 24/25 ("B" in Table 5-1). The

Table 5-1. Simple Bayesian Analysis Generalized

	Hypothesis	
	1	2
Prior probability	A	B = 1 – A
Conditional probability	C	D
Joint probability	E = AC	F = BD
Posterior probability	G = E ÷ (E + F)	H = F ÷ (E + F)

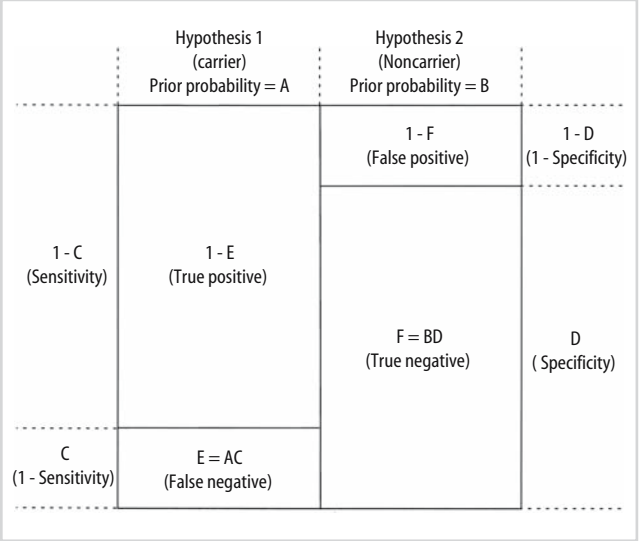


Figure 5-3. Schematic representation of the generalized Bayesian analysis shown in Table 5-1, for the case of a negative carrier test. The small rectangles represent true-positive, false-positive, true-negative, and false-negative rates for a particular consultand; that is, the prior probabilities are influenced by factors such as family history or signs and symptoms, and the sensitivity and specificity of the test are influenced by factors such as ethnicity. For a negative carrier test, the posterior carrier probability (1 minus the negative predictive value) is the false-negative rate divided by the sum of the false- and true-negative rates, or E ÷ (E + F).

ACMG screening guidelines panel of 23 mutations detects 94% of CF mutations in Ashkenazi Jewish populations,[10–12] so the conditional probability of a negative test, under the hypothesis that he is a carrier, is 6% = 3/50 ("C" in Table 5-1). Under the hypothesis that he is a noncarrier, the conditional probability of a negative test approximates 1 ("D" in Table 5-1). (This is generally the case in genetic testing, since noncarriers by definition lack mutations in the relevant disease gene and, hence, unless there are technical problems, essentially always should test negative.) The Bayesian analysis table for this example is shown in Table 5-2. The joint probabilities are the products of the prior and conditional probabilities ("E" and "F" in Table 5-1), and the posterior probabilities ("G" and "H" in Table 5-1) derive from each joint probability divided by the sum of the joint probabilities. The husband's posterior carrier risk after the negative test result is 1/401 (Table 5-2).

What is the risk that the fetus of the mother (consultand) in Figure 5-2 and the father from Table 5-2 is affected with CF? Prior to testing, the risk was the prior probability that the mother was a carrier (2/3), multiplied by the prior probability that the father was a carrier (1/25), multiplied by the probability that the fetus would inherit two disease alleles (1/4), or $2/3 \times 1/25 \times 1/4 = 1/150$. After testing, the risk is the posterior probability that the mother is a carrier (1/6), multiplied by the posterior probability that the father is a carrier (1/401), multiplied by the probability that the fetus would inherit two disease alleles (1/4), or $1/6 \times 1/401 \times 1/4 \cong 1/9600$.

Often, testing is performed on additional family members, and genetic risks need to be modified accordingly. In the example above, testing of both parents of the mother (consultand) would affect her carrier risk calculations. Detection of mutations in both parents using the same mutation test panel would essentially rule out carrier status for the mother, since we would then know that the sensitivity of the test for the mutations she is at risk of carrying is essentially 100%. Alternatively, if the test results for the mother's parents are positive for only one of her parents (for example, her father) and negative for the other parent (her mother), then the sensitivity of the test for the mutations she is at risk of carrying is essentially 50%. The Bayesian analysis for the mother, modified from Figure 5-2c, is shown in Table 5-3a. The conditional probability of a negative test under the hypothesis that she is a carrier has changed from 1/10 to 1/2, which increases the posterior probability that she is a carrier to 1/2. Taken together with her husband's carrier risk of 1/401 (Table 5-2), the risk that

Table 5-2. Bayesian Analysis for an Ashkenazi Jewish Individual Without a Family History of CF Who Tests Negative for the ACMG Screening Guidelines Panel of 23 *CFTR* Mutations

	Hypothesis	
	Carrier	Noncarrier
Prior probability	1/25	24/25
Conditional probability (of negative test result)	3/50	1
Joint probability	3/1250	24/25
Posterior probability	(3/1250) ÷ (3/1250 + 24/25) = 1/401	(24/25) ÷ (3/1250 + 24/25) = 400/401

Table 5-3a. Bayesian Analysis for the Consultand in Figure 5-2a After Testing of the Parents (see text)

	Hypothesis	
	Carrier	Noncarrier
Prior probability	2/3	1/3
Conditional probability (of negative test result)	1/2	1
Joint probability	1/3	1/3
Posterior probability	(1/3) ÷ (1/3 + 1/3) = 1/2	(1/3) ÷ (1/3 + 1/3) = 1/2

Table 5-3b. Alternative Bayesian Analysis for the Consultand in Figure 5-2a After Testing of the Parents (see text)

	Hypothesis		
	Carrier		
	Carrier with Paternal (detectable) Mutation	Carrier with Maternal (undetectable) Mutation	Noncarrier
Prior probability	1/3	1/3	1/3
Conditional probability (of negative test result)	0	1	1
Joint probability	0	1/3	1/3
Posterior probability	0	1/2	1/2

the fetus is affected with CF can be modified to $1/2 \times 1/401 \times 1/4 \cong 1/3200$.

Another way of conceptualizing the Bayesian analysis described above is to separate the carrier hypothesis into two subhypotheses, as shown in Table 5-3b. (See also "Bayesian Analysis with More Than Two Hypotheses," below.) The two subhypotheses are (1) that the consultand is a carrier with a paternal (detectable) mutation and (2) that she is a carrier with a maternal (undetectable) mutation. The prior probability of each hypothesis is 1/3, that is, half of 2/3. The conditional probability of a negative test result, under the subhypothesis that she is a carrier of a detectable paternal mutation, is 0. The conditional probability of a negative test result, under the subhypothesis that she is a carrier of an undetectable maternal mutation, is 1. As in the generalized Bayesian analysis shown in Table 5-1, the joint probability for each hypothesis is the product of the prior and conditional probabilities for that hypothesis, and the posterior probability for each hypothesis is the joint probability for that hypothesis divided by the sum of all the joint probabilities. The posterior probability that the consultand has a detectable paternal mutation is 0, and the posterior probability that she has an undetectable maternal mutation is 1/2 (Table 5-3b).

Simple Bayesian Analyses Generalized: Affected Versus Unaffected

Another common application of Bayesian analysis in molecular pathology is to calculate the risk that a patient is affected with a particular disease after a negative test result. Again, the need to calculate risk in this scenario stems from the fact that the sensitivities of many genetic tests are less than 100%. Hypothesis 1 (in Table 5-1 and Figure 5-3) in this scenario is that the patient is affected, and Hypothesis 2 is that the patient is unaffected. The prior probability ("A" in Table 5-1 and Figure 5-3) usually derives mostly from signs and symptoms but also may depend on aspects of the patient's history, including family history in diseases with a genetic component. As in the CF example, above, the conditional probabilities ("C" and "D" in Table 5-1 and Figure 5-3) are 1 minus the test sensitivity, and the test specificity, respectively. The remainder of the analysis is accomplished by calculation, with the posterior probabilities ("G" and "H" in Table 5-1) representing 1 minus the negative predictive value, and the negative predictive value, respectively.

For example, suppose that a child with clinically typical spinal muscular atrophy type III (type III SMA; Kugelberg-Welander disease; OMIM #253400) tests negative for the homozygous deletion of the *SMN1* gene found in most affected individuals. What is the probability that the child is affected with *SMN1*-linked SMA? The Bayesian analysis for this scenario is shown in Table 5-4a. Wirth et al. found that 17 of 131 individuals with clinically typical type III SMA lacked mutations in both *SMN1* alleles (and therefore were considered to have diseases unrelated to *SMN1*);[13] hence, the prior probability that the child is affected with *SMN1*-linked type III SMA is 114/131, or 0.87. Approximately 6% of individuals with *SMN1*-linked type III SMA have a deletion of one *SMN1* allele and a subtle mutation, undetectable by simple polymerase chain reaction (PCR) testing for a homozygous deletion, in the other *SMN1* allele;[14] hence, the conditional probability of a negative test result under the hypothesis that the child is affected is 6/100 or 0.06. Homozygous deletions of *SMN1*, when present, are highly specific for *SMN1*-linked SMA; hence, the conditional probability of a negative test result under the hypothesis that the child is unaffected with *SMN1*-linked SMA approximates 1. Following the simple calculation rules in Table 5-1, the posterior probability that the child is affected with *SMN1*-linked type III SMA is approximately 0.29 (Table 5-4a).

Suppose that *SMN1* dosage analysis is performed on the child's DNA (i.e., the SMA carrier test), and the result is that the child has one copy of the *SMN1* gene. What is the probability that he or she is affected with *SMN1*-linked SMA? The Bayesian analysis for this scenario is shown in Table 5-4b. Again, the prior probability that the child is affected with *SMN1*-linked type III SMA is 0.87. Because approxi-

Table 5-4a. Bayesian Analysis for a Child with Clinically Typical Type III SMA Who Tests Negative for Homozygous Deletions of the *SMN1* Gene

	Hypothesis	
	Affected	Unaffected
Prior Probability	0.87	0.13
Conditional Probability (of negative test result)	0.06	~1
Joint Probability	0.052	0.13
Posterior Probability	0.29	0.71

Table 5-4b. Bayesian Analysis for a Child with Clinically Typical Type III SMA Who Has One Copy of the *SMN1* Gene by Dosage Analysis

	Hypothesis	
	Affected	Unaffected
Prior probability	0.87	0.13
Conditional probability (of 1-copy test result)	0.06	0.026
Joint probability	0.052	0.0034
Posterior probability	0.94	0.06

Table 5-4c. Bayesian Analysis for a Child with Clinically Typical Type III SMA Who Has 2 Copies of the *SMN1* Gene by Dosage Analysis

	Hypothesis	
	Affected	Unaffected
Prior probability	0.87	0.13
Conditional probability (of 2-copy test result)	0.0009	0.9
Joint probability	0.00078	0.12
Posterior probability	0.006	0.994

mately 6% of individuals with *SMN1*-linked type III SMA have a deletion of one *SMN1* allele and a subtle mutation in the other *SMN1* allele that is detectable as a single copy by dosage analysis,[14] the conditional probability of a single-copy test result under the hypothesis that the child is affected is again 0.06. However, the carrier frequency for SMA in the general population is approximately 1/38;[14] hence, in this scenario the conditional probability of a single-copy test result under the hypothesis that the child is unaffected with *SMN1*-linked SMA is 1/38 or 0.026. Following the simple calculation rules in Table 5-1, the posterior probability that the child is affected with *SMN1*-linked type III SMA is approximately 0.94 (Table 5-4b).

Suppose instead that the result of the *SMN1* dosage analysis is that the child has two copies of the *SMN1* gene. What is the probability that the child is affected with *SMN1*-linked SMA? The Bayesian analysis for this scenario is shown in Table 5-4c. Again, the prior probability that the child is affected with *SMN1*-linked type III SMA is 0.87. Only approximately 9 in 10,000 individuals with *SMN1*-linked type III SMA would be expected to have two subtle, nondeletion mutations, detectable as two gene copies by dosage analysis;[14] hence, the conditional probability of a two-copy test result under the hypothesis that the child is affected is approximately 0.0009. Because more than 7% of unaffected individuals have three copies of the *SMN1* gene, and approximately 2.5% of unaffected individuals have one copy of the *SMN1* gene, for a total of 9.5% of unaffected individuals without two copies of *SMN1*,[14] the conditional probability of a 2-copy test result under the hypothesis that the child is unaffected with *SMN1*-linked SMA is 90.5/100, or approximately 0.9. Following the simple calculation rules in Table 5-1, the posterior probability that the child is affected with *SMN1*-linked type III SMA is only approximately 0.006 (Table 5-4c).

Profiling by proteomics, RNA microarrays, or analysis of single-nucleotide polymorphisms (SNPs), or some combination of these, is likely to play an important role in molecular pathology in the future, and clinical test results will be reported, in many cases, as probabilities or relative risks. For example, suppose that a consultand has a 20% lifetime risk of developing a particular disease (based on family history, physical examination, or clinical laboratory test results, or a combination of these) and that his or her proteomic profile is 16 times more common in those who go on to develop the disease than in those who do not. What is his or her lifetime risk of developing the disease? The Bayesian analysis for this scenario is shown in Table 5-5. Hypothesis 1 (from Table 5-1) is that the consultand will develop the disease, and Hypothesis 2 is that the consultand will not develop the disease. The prior probabilities are 0.2 and 0.8 for Hypotheses 1 and 2, respectively. Because the conditional probability of the proteomic profiling result is 16 times more likely in those who develop the disease than in those who do not, the conditional probabilities ("C" and "D" in Table 5-1) are 16 and 1, respectively. Following the simple calculation rules in Table 5-1, the posterior probability that the consultand will develop the disease is 0.8 (Table 5-5).

Table 5-5. Bayesian Analysis for a Consultand with a 20% Lifetime Risk of Developing a Disease and a Proteomic Profile 16 Times More Common in Those Who Develop the Disease Than in Those Who Do Not

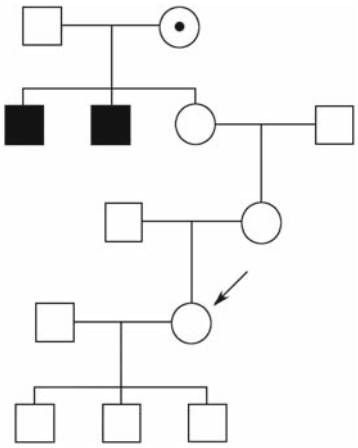
	Hypothesis	
	Affected Eventually	Never Affected
Prior probability	0.2	0.8
Conditional probability (of profiling result)	16	1
Joint probability	3.2	0.8
Posterior probability	0.8	0.2

Note that because posterior probabilities are normalized joint probabilities, the absolute values of the conditional probabilities are unimportant, as long as the ratio (i.e., the odds ratio) between them is correct. This is also true of prior probabilities. For example, in the scenario above, prior probabilities of 1 and 4 can be substituted for 0.2 and 0.8 and the same answer is obtained. Likewise, in the first example of this chapter (Figure 5-1a), prior probabilities of 1 and 1 can be substituted for 1/2 and 1/2, and conditional probabilities of 1 and 8 can be substituted for 1/8 and 1, and the same answer is obtained. Hence, relative risks are easily incorporated into Bayesian analyses.

Bayesian Analyses with More Than One Conditional Probability

Often there is more than one test result, or more than one set of pedigree information, or both, that can be incorporated as conditional probabilities in a single Bayesian analysis. For example, consider the pedigree in Figure 5-4a, in which the two maternal great uncles of the consultand were affected with Duchenne muscular dystrophy (DMD; OMIM #310200), a severe X-linked recessive disease caused by mutations in the *DMD* gene (OMIM #300377). The consultand's maternal grandmother's carrier risk was 1/2, her mother's carrier risk was 1/4, and therefore the consultand's prior carrier risk is 1/8. Suppose that her carrier testing is negative using a highly specific test (an analysis for heterozygous deletions in the *DMD* gene) that detects 2/3 of carriers. Suppose also that her serum creatine phosphokinase (CPK), which is elevated in two-thirds of carriers, is within normal limits. Taking into account her prior probability of 1/8, her normal molecular and CPK test results, and, in addition, her three normal sons, what is the probability that she is a carrier?

The Bayesian analysis for this scenario is shown in Figure 5-4b. Each conditional probability is given its own line. Because the genetic test detects 2/3 of carriers and is highly specific, the conditional probabilities of a negative genetic test result under the hypotheses that she is a carrier and noncarrier are 1/3 and 1, respectively. Because serum CPK is elevated in 2/3 of carriers, the conditional



a

	Hypothesis	
	Carrier	Non-carrier
Prior Probability	1/8	7/8
Conditional probability (of negative genetic test result)	1/3	1
Conditional Probability (of normal CPK result)	1/3	19/20
Conditional Probability (of three normal sons)	1/8	1
Joint Probability	1/576	133/160
Posterior Probability	0.002	0.998

b

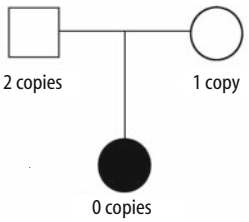
Figure 5-4. (a) Pedigree of a family with individuals affected with DMD (see text). Consultand is indicated by an arrow. (b) Bayesian analysis for the consultand in Figure 5-4a, taking into account her normal carrier test result, her normal CPK test result, and her three normal sons.

probability of a normal serum CPK for the hypothesis that she is a carrier is 1/3. Because 5% of noncarrier women have an abnormal serum CPK (i.e., the normal range is defined as comprising 95% of normal individuals), the conditional probability of a normal serum CPK under the hypothesis that she is a noncarrier is 95% or 19/20. Finally, as in Figure 5-1b, the conditional probabilities of three normal sons under the hypotheses that she is a carrier and noncarrier are 1/8 and 1, respectively. The joint probabilities for each hypothesis are the products of the prior probability, and all conditional probabilities, for each hypothesis (Figure 5-4b). Calculation of posterior probabilities then proceeds exactly as in Table 5-1. In this scenario, taking into account her normal test results and her three normal sons, the consultand's carrier risk is lowered from 1/8 to 0.002, or approximately 1/500.

Bayesian Analyses with More Than Two Hypotheses

In some Bayesian analyses, more than two hypotheses must be considered. For example, consider the pedigree in Figure 5-5a, in which a child with clinically typical type I SMA (type I SMA; Werdnig-Hoffman disease; OMIM #253300) lacks both copies of the *SMN1* gene. By dosage analysis, the child's (unaffected) mother has one copy of the *SMN1* gene and therefore carries one copy of the *SMN1* gene on one chromosome 5, and zero copies of the *SMN1* gene on the other chromosome 5, called the "1 + 0" genotype. However, the child's (unaffected) father has two copies of the *SMN1* gene and therefore could have one of three possible genotypes: (1) two copies of the *SMN1* gene on one chromosome 5 and zero copies of the *SMN1* gene on the other chromosome 5 (the "2 + 0" genotype), (2) one copy of the *SMN1* gene on one chromosome 5 and a subtle mutation in the *SMN1* gene on the other chromosome 5 (the "1 + 1^D" genotype, where "1^D" stands for a "1-copy-disease" allele), or (3) one copy of the *SMN1* gene on each chromosome 5 (the "1 + 1" noncarrier genotype), in which case he passed a de novo deletion of the *SMN1* gene to his affected child. Because the relative frequencies of the various *SMN1* alleles and genotypes in the general population are known,[14] as well as the paternal and maternal de novo deletion rates ($\mu_p = 2.11 \times 10^{-4}$ and $\mu_m = 4.15 \times 10^{-5}$, respectively), the probability that the father is a carrier can be calculated, which obviously has important implications for recurrence risk.

The Bayesian analysis for the father's carrier risk is shown in Figure 5-5b. There are three hypotheses for the father's genotype: 2 + 0, 1 + 1, and 1 + 1^D. The prior probabilities are the relative population frequencies for these genotypes.[14] The conditional probabilities are the probabilities that the father passes a 0-copy allele to his child under each hypothesis. For the 2 + 0 genotype, the conditional probability of passing a 0-copy allele is 0.5, whereas for the 1 + 1 and 1 + 1^D genotypes, the conditional probability of passing a 0-copy allele is the de novo deletion rate of μ_p. As in the generalized Bayesian analysis shown in Table 5-1, the joint probability for each hypothesis is the



a

	Hypothesis (Father's Genotype)		
	2 + 0	1 + 1	1 + 1^D
Prior Probability (relative probability)	1.00×10^{-3}	9.00×10^{-1}	4.58×10^{-4}
Conditional Probability (of passing a 0-copy allele)	0.5	μ_p	μ_p
Joint Probability	5.0×10^{-4}	1.9×10^{-4}	9.7×10^{-8}
Posterior Probability	0.72	0.28	0.00014

b

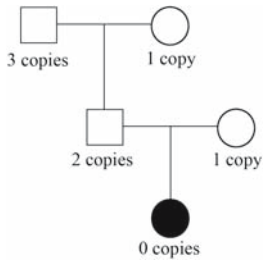
Figure 5-5. (a) Pedigree of a family with an individual affected with type I SMA, with the SMA carrier test results indicated below each individual (see text). (b) Bayesian analysis for carrier risk of the father of the affected child in Figure 5-5a. μ_p, paternal de novo mutation rate.

product of the prior and conditional probabilities for that hypothesis, and the posterior probability for each hypothesis is the joint probability for that hypothesis divided by the sum of all the joint probabilities. The father's carrier risk is the sum of the posterior probabilities of the first (2 + 0) and third (1 + 1^D) columns, or approximately 0.72. The third column contributes little to the carrier risk because the frequency of the 1 + 1^D genotype is low and the conditional probability of a de novo deletion is also low. In contrast, although the frequency of the 2 + 0 genotype is much lower than that of the 1 + 1 genotype, this is counterbalanced by the higher conditional probability of passing a 0-copy allele under the former hypothesis.

Suppose that the father's parents, the paternal grandfather and grandmother of the affected child, are tested and found to have three copies and one copy of the *SMN1* gene, respectively (Figure 5-6a). What is the father's carrier risk? The Bayesian analysis for this scenario is shown in Figure 5-6b. Again, there are three hypotheses for the father's genotype: 2 + 0, 1 + 1, and 1 + 1^D. However, in this scenario, the father's prior probabilities derive from the prior and conditional probabilities of his parents. Because the grandfather has three copies of the *SMN1* gene, his genotype is either 2 + 1 (columns A and C) or 2 + 1^D (columns B and D), and his prior probabilities are the relative population frequencies for these genotypes.[14] Because the (unaffected) grandmother has one copy of the *SMN1* gene, her genotype is 1 + 0, and her prior probability is the relative population frequency of the 1 + 0 genotype for type I SMA in the general population, which is the carrier frequency of 1/38 (2.50×10^{-2}).[14] (Note that because the grandmother must be 1 + 0, simply a prior probability of 1 could be used; as noted above, the absolute values of the conditional probabilities are unimportant, as long as the ratio between them is correct.) The four columns (A through D) show the four possible permutations of grandparental genotypes (prior probabilities) with passage of particular alleles to the father (conditional probabilities) so that he would have a 2-copy SMA carrier test result. Under the hypothesis that the father has a 2 + 0 genotype, he could have inherited a 2-copy "allele" (two copies of *SMN1* on one chromosome 5) from the grandfather (2 + 1) at a probability of 0.5 and a 0-copy allele from the grandmother (1 + 0) at a probability of 0.5 (column A), or he could have inherited a 2-copy allele from the grandfather (2 + 1^D) at a probability of 0.5 and a 0-copy allele from the grandmother (1 + 0) at a probability of 0.5 (column B). Under the hypothesis that the father has a 1 + 1 genotype, he could have inherited a 1-copy allele from the grandfather (2 + 1) at a probability of 0.5 and a 1-copy allele from the grandmother (1 + 0) at a probability of 0.5 (column C). Under the hypothesis that the father has a 1 + 1^D genotype, he could have inherited a 1^D allele from the grandfather (2 + 1^D) at a probability of 0.5 and a 1-copy allele from the grandmother (1 + 0) at a probability of 0.5 (column D).

The father's prior probabilities are the products of the prior and conditional probabilities for the grandparents for each column or permutation. Under the hypothesis that the father has a 2 + 0 genotype, the conditional probability of passing a 0-copy allele to his child is 0.5 (columns A and B), whereas under the hypothesis that the father has a 1 + 1 or 1 + 1^D genotype, the conditional probability of passing a 0-copy allele to his child is the de novo deletion rate of μ_p (columns C and D). As in the generalized Bayesian analysis shown in Table 5-1, the joint probability for each column is the product of the prior and conditional probabilities for that column, and the posterior probability for each column is the joint probability for that column divided by the sum of all the joint probabilities. The father's carrier risk is the sum of the posterior probabilities of columns A (2 + 0), B (2 + 0), and D (1 + 1^D), or approximately 0.999. The father's increased carrier risk in this scenario derives almost entirely from the probability that he has the 2 + 0 genotype; this is unsurprising since the grandfather's 3-copy test result demonstrates the presence of a 2-copy allele in the family. (Note that because the grandmother's prior and conditional probabilities are the same in every column, excluding her data from the analysis will not change the result.)

Suppose instead that the father's parents, the paternal grandfather and grandmother of the affected child, are tested and each is found to have two copies of the *SMN1* gene (Figure 5-7a). What is the father's carrier risk? The Bayesian analysis for this scenario is shown in Figure 5-7b. Again, there are three hypotheses for the father's genotype:

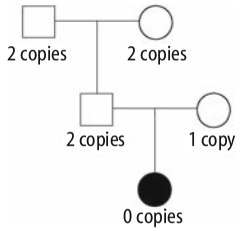


a

	Hypothesis (Father's Genotype)			
	2 + 0		1 + 1	1 + 1^D
	A	B	C	D
Grandfather's Genotype	2 + 1	2 + 1^D	2 + 1	2 + 1^D
Prior probability (relative probability)	7.18×10^{-2}	1.15×10^{-5}	7.18×10^{-2}	1.15×10^{-5}
Conditional probability (of passing a 2-copy allele)	0.5	0.5	--	--
Conditional probability (of passing a 1 or 1^D allele)	--	--	0.5	0.5
Grandmother's Genotype	1 + 0	1 + 0	1 + 0	1 + 0
Prior probability	2.50×10^{-2}	2.50×10^{-2}	2.50×10^{-2}	2.50×10^{-2}
Conditional probability (of passing a 0-copy allele)	0.5	0.5	--	--
Conditional probability (of passing a 1-copy allele)	--	--	0.5	0.5
Father's Genotype	2 + 0		1 + 1	1 + 1^D
Conditional probability (of passing a 0-copy allele)	0.5 0.5		μ_p	μ_p
Joint Probability	2.24×10^{-4}	3.59×10^{-8}	9.47×10^{-8}	1.52×10^{-11}
Posterior Probability	0.999	0.00016	0.00042	0.000000068

b

Figure 5-6. (a) Pedigree of a family with an individual affected with type I SMA with the SMA carrier test results indicated below each individual (see text). (b) Bayesian analysis for carrier risks of the father of the affected child in Figure 5-6a.



a

	Hypothesis (Father's Genotype)													
	2 + 0						1 + 1				1 + 1^{D}			
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
Grandfather's Genotype	2 + 0	2 + 0	2 + 0	2 + 0	1 + 1	1 + 1^{D}	1 + 1	1 + 1	1 + 1^{D}	1 + 1^{D}	1 + 1	1 + 1^{D}	1 + 1^{D}	1 + 1^{D}
Prior Probability	1.0 x 10^{-3}	1.0 x 10^{-3}	1.0 x 10^{-3}	1.0 x 10^{-3}	0.90	4.6 x 10^{-4}	0.90	0.90	4.6 x 10^{-4}	4.6 x 10^{-4}	0.90	4.6 x 10^{-4}	4.6 x 10^{-4}	4.6 x 10^{-4}
Conditional Probability (of passing a 2-copy allele)	0.5	--	0.5	0.5	--	--	--	--	--	--	--	--	--	--
Conditional Probability (of passing a 0-copy allele)	--	0.5	--	--	2.1 x 10^{-4}	2.1 x 10^{-4}	--	--	--	--	--	--	--	--
Conditional Probability (of passing a 1-copy allele)	--	--	--	--	--	--	1	1	0.5	0.5	1	--	--	0.5
Conditional Probability (of passing a 1^{D} allele)	--	--	--	--	--	--	--	--	--	--	--	0.5	0.5	--
Grandmother's Genotype	2 + 0	2 + 0	1 + 1	1 + 1^{D}	2 + 0	2 + 0	1 + 1	1 + 1^{D}	1 + 1	1 + 1^{D}	1 + 1^{D}	1 + 1	1 + 1^{D}	1 + 1^{D}
Prior Probability	1.0 x 10^{-3}	1.0 x 10^{-3}	0.90	4.6 x 10^{-4}	1.0 x 10^{-3}	1.0 x 10^{-3}	0.90	4.6 x 10^{-4}	0.90	4.6 x 10^{-4}	4.6 x 10^{-4}	0.90	4.6 x 10^{-4}	4.6 x 10^{-4}
Conditional Probability (of passing a 2-copy allele)	--	0.5	--	--	0.5	0.5	--	--	--	--	--	--	--	--
Conditional Probability (of passing a 0-copy allele)	0.5	--	4.2 x 10^{-5}	4.2 x 10^{-5}	--	--	--	--	--	--	--	--	--	--
Conditional Probability (of passing a 1-copy allele)	--	--	--	--	--	--	1	0.5	1	0.5	--	1	0.5	--
Conditional Probability (of passing a 1^{D} allele)	--	--	--	--	--	--	--	--	--	--	0.5	--	--	0.5
Father's Genotype	2 + 0						1 + 1				1 + 1^{D}			
Conditional Probability (of passing a 0-copy allele)	0.5	0.5	0.5	0.5	0.5	0.5	2.1 x 10^{-4}	2.1 x 10^{-4}	2.1 x 10^{-4}	2.1 x 10^{-4}	2.1 x 10^{-4}	2.1 x 10^{-4}	2.1 x 10^{-4}	2.1 x 10^{-4}
Joint Probability	1.3 x 10^{-7}	1.3 x 10^{-7}	9.3 x 10^{-9}	4.8 x 10^{-12}	4.8 x 10^{-8}	2.4 x 10^{-11}	1.7 x 10^{-4}	4.3 x 10^{-8}	4.3 x 10^{-8}	1.1 x 10^{-11}	4.3 x 10^{-8}	4.3 x 10^{-8}	1.1 x 10^{-11}	1.1 x 10^{-11}
Posterior Probability	7.3 x 10^{-4}	7.3 x 10^{-4}	5.4 x 10^{-5}	2.8 x 10^{-8}	2.8 x 10^{-4}	1.4 x 10^{-7}	0.997	2.5 x 10^{-4}	2.5 x 10^{-4}	6.5 x 10^{-8}	2.5 x 10^{-4}	2.5 x 10^{-4}	6.5 x 10^{-8}	6.5 x 10^{-8}

b

Grandfather's Genotype (2 copies and asymptomatic)	1+1				2+0			1+1^{D}				
Prior Probability	0.9				0.001			0.00046				
Grandmother's Genotype (2 copies and asymptomatic)	1+1	2+0	1+1^{D}		1+1	2+0	1+1^{D}	1+1		2+0	1+1^{D}	
Prior Probability	0.9	0.001	0.00046		0.9	0.001	0.00046	0.9		0.001	0.00046	
Father's Genotype (2 copies and asymptomatic)	1+1 NC	2+0 C	1+1 NC	1+1^{D} C	2+0 C	2+0 C	2+0 C	1+1 NC	1+1^{D} C	2+0 C	1+1 NC	1+1^{D} C
Conditional Probability (of receiving alleles)	1	0.5 μ_p	0.5	0.5	0.5 μ_m	0.5	0.5 μ_m	0.5	0.5	0.5 μ_p	0.25	0.5
Conditional Probability (of passing a 0-copy allele)	μ_p	0.5	μ_p	μ_p	0.5	0.5	0.5	μ_p	μ_p	0.5	μ_p	μ_p
Joint Probability	1.71x10^{-4}	4.75x10^{-8}	4.35x10^{-8}	4.35x10^{-9}	9.34x10^{-9}	2.5x10^{-7}	4.75x10^{-12}	4.35x10^{-8}	4.35x10^{-8}	2.42x10^{-11}	1.11x10^{-11}	2.22x10^{-11}
Posterior Probability	0.997	2.77x10^{-4}	2.54x10^{-4}	2.54x10^{-4}	5.45x10^{-5}	0.00146	2.77x10^{-7}	2.54x10^{-4}	2.54x10^{-4}	1.41x10^{-7}	6.47x10^{-8}	1.29x10^{-7}
Column	A	B	C	D	E	F	G	H	I	J	K	L

c

Figure 5-7. (a) Pedigree of a family with an individual affected with type I SMA (see text). (b) Bayesian analysis for the father of the affected child in Figure 5-7a. (In the interests of space, only two significant digits are shown.) (c) Alternative Bayesian analysis for the father of the affected child in Figure 5-7a. NC, noncarrier; C, carrier.

2 + 0, 1 + 1, and 1 + 1^D. However, in this scenario, the number of possible permutations of grandparental genotypes (prior probabilities) with passage of particular alleles to the father (conditional probabilities) is dramatically increased. This is because each grandparent could have a 2 + 0, 1 + 1, or 1 + 1^D genotype, and the father could have received a 2-copy allele, a 0-copy allele, a 1-copy allele, or a 1^D allele from either grandparent, in most cases by direct Mendelian inheritance and in some cases from de novo deletions. The organization of the Bayesian analysis in Figure 5-7b is guided by the possible genotypes of the father, which determine the grandparental genotype permutations that need to be considered. Under the hypothesis that the father has the 2 + 0 genotype, he could have received a 2-copy allele from one (2 + 0) grandparent and a 0-copy allele from the other (2 + 0) grandparent, both by direct inheritance (columns A and B), or he could have received a 2-copy allele from one (2 + 0) grandparent by direct inheritance and a de novo deletion allele from the other (1 + 1 or 1 + 1^D) grandparent (columns C, D, E, and F). Under the hypothesis that the father has the 1 + 1 genotype, he must have received a 1-copy allele from each (1 + 1 or 1 + 1^D) grandparent (columns G, H, I, and J). Under the hypothesis that the father has the 1 + 1^D genotype, he must have received a 1-copy allele from one (1 + 1 or 1 + 1^D) grandparent and a 1^D allele from the other (1 + 1 or 1 + 1^D) grandparent (columns K, L, M, and N).

More specifically, under the hypothesis that the father has the 2 + 0 genotype, column A shows the prior probability that the grandfather is 2 + 0 (1.00×10^{-3}), the conditional probability that he passes a 2-copy allele to the father (0.5), the prior probability that the grandmother is 2 + 0 (1.00×10^{-3}), and the conditional probability that she passes a 0-copy allele to the father (0.5). Under the hypothesis that the father has the 1 + 1 genotype, column G shows the prior probability that the grandfather has a 1 + 1 genotype (0.90), the conditional probability that he passes a 1-copy allele to the father (1), the prior probability that the grandmother has a 1 + 1 genotype (0.90), and the conditional probability that she passes a 1-copy allele to the father (1). Under the hypothesis that the father has the 1 + 1^D genotype, column K shows the prior probability that the grandfather has a 1 + 1 genotype (0.90), the conditional probability that he passes a 1-copy allele to the father (1), the prior probability that the grandmother has a 1 + 1^D genotype (4.58×10^{-4}), and the conditional probability that she passes a 1^D allele to the father (0.5).

Again, the father's prior probabilities are the products of the prior and conditional probabilities for the grandparents for each column or permutation. Under the hypothesis that the father has a 2 + 0 genotype, the conditional probability of passing a 0-copy allele to his child is 0.5 (columns A through F), whereas under the hypothesis that the father has a 1 + 1 or 1 + 1^D genotype, the conditional probability of passing a 0-copy allele to his child is the paternal de novo deletion rate of μ_p (columns G through L). As in the generalized Bayesian analysis shown in Table 5-1, the joint probability for each column is the product of the prior and conditional probabilities for that column, and the posterior probability for each column is the joint probability for that column divided by the sum of all the joint probabilities. The father's carrier risk is the sum of the posterior probabilities of columns A through F (2 + 0), and K and L (1 + 1^D), or approximately 1/400. Relative to the previous scenario (Figure 5-6), in which the father also had two copies of *SMN1* but the grandparents had different copy numbers, the father's dramatically decreased carrier risk in this scenario derives from the much lower probability that a 2-copy allele is present in his family, and illustrates the importance of integrating all available genetic testing information into risk assessment calculations.

An alternative organization of the Bayesian analysis shown in Figure 5-7b is shown in Figure 5-7c and is guided by the three hypotheses for the grandparental genotypes: 1 + 1, 2 + 0, and 1 + 1^D. For example, under the hypothesis that both of the grandparents have a 1 + 1 genotype, column A shows the prior probabilities that the grandfather has a 1 + 1 genotype (0.9) and that the grandmother has a 1 + 1 genotype (0.9), the conditional probability that the father received 1-copy alleles from both of the grandparents (1), and the conditional probability that the father passed a 0-copy allele to the affected child (by de novo deletion, μ_p). Under the hypothesis that the grandfather has a 1 + 1 genotype and that the grandmother has a 2 + 0 genotype, column B shows the prior probabilities that the grandfather has a 1 + 1 genotype (0.90) and that the grandmother has a 2 + 0 genotype (0.001), the conditional probability that the father received a 2-copy allele from one of the grandparents (the grandmother in this case) (0.5) and a 0-copy allele from the other grandparent (the grandfather in this case by de novo deletion, μ_p), and the conditional probability that the father passed a 0-copy allele to the affected child (0.5). Under the hypothesis that the grandfather has the 1 + 1 genotype and that the grandmother has the 1 + 1^D genotype, column C shows the prior probabilities that the grandfather has a 1 + 1 genotype (0.90) and that the grandmother has a 1 + 1^D genotype (0.00046), the conditional probability that the father received 1-copy alleles from both grandparents (0.5), and the conditional probability that the father passed a 0-copy allele to the affected child (by de novo deletion, μ_p). Under the hypothesis that the grandfather has the 1 + 1 genotype and that the grandmother has the 1 + 1^D genotype, column D shows the prior probabilities that the grandfather has a 1 + 1 genotype (0.90) and that the grandmother has a 1 + 1^D genotype (0.00046), the conditional probability that the father received a 1^D allele from one of the grandparents (the grandmother in this case) and a 1-copy allele from the other grandparent (the grandfather in this case) (0.5), and the conditional probability that the father passed a 0-copy allele to the affected child (by de novo deletion, μ_p). The father's carrier risk is the sum of the posterior probabilities of columns B, D through G, I, J, and L, or approximately 1/400.

In both approaches (Figures 5-7b and 5-7c), the use of one comprehensive Bayesian analysis table incorporating all necessary information allows simultaneous calculations of the carrier risks of the father, grandfather, and grandmother. Such a comprehensive approach is necessary because the 2-copy test results for the grandparents influence the carrier risk of the father, and the 2-copy test result for the father influences the carrier risks of the grandparents. Using Figure 5-7b, the posterior carrier risk of the grandfather is the sum of the posterior probabilities of columns A through D, F, I, J, and L through N, or approximately 0.0020 (1/500), and the carrier risk of the grandmother is the sum of the posterior probabilities of columns A, B, D through F, H, J, K, M, and N, or approximately 0.0022 (1/450). The posterior probability that all three of them are carriers is the sum of the posterior probabilities of columns A, B, D, F, M, and N, or approximately 0.0015 (1/600). Using Figure 5-7c, the carrier risk of the grandfather is the sum of the posterior probabilities of columns E through L, or approximately 0.0020 (1/500), and the carrier risk of the grandmother is the sum of the posterior probabilities of columns B through D, F, G, and J through L, or approximately 0.0022 (1/450). The posterior probability that all three of them are carriers is the sum of the posterior probabilities of columns F, G, J, and L, or approximately 0.0015 (1/600).

Concluding Remarks

Bayesian analysis plays a central role in genetic risk assessment, and those who offer genetic testing should be proficient. Genetic risk should be assessed as accurately as possible, using all available information at a particular point in time, from the pedigree, from laboratory testing, or from both. Although the technologies for genetic testing will continue to change, Bayesian analysis and genetic risk assessment will remain fundamental aspects of genetic testing and genetic counseling.

References

1. Bayes T. An essay towards solving a problem in the doctrine of chances. *Biometrika.* 1958;45:296–315.
2. Young I. *Introduction to Risk Calculation in Genetic Counseling.* Oxford: Oxford University Press; 1999.
3. Bridge P. *The Calculation of Genetic Risks: Worked Examples in DNA Diagnostics.* Baltimore: Johns Hopkins University Press; 1997.
4. Ogino S, Leonard DGB, Rennert H, Ewens WJ, Wilson RB. Genetic risk assessment in carrier testing for spinal muscular atrophy. *Am J Med Genet.* 2002;110:301–307.
5. Ogino S, Wilson RB. Genetic testing and risk assessment for spinal muscular atrophy (SMA). *Hum Genet.* 2002;111:477–500.
6. Ogino S, Wilson RB. Bayesian analysis and risk assessment in genetic counseling and testing. *J Mol Diagn.* 2004;6:1–9.
7. Ogino S, Wilson RB, Gold B, Hawley P, Grody WW. Bayesian analysis for cystic fibrosis risks in prenatal and carrier screening. *Genet Med.* 2004;6:439–449.
8. Ogino S, Wilson RB, Grody WW. Bayesian risk assessment for autosomal recessive diseases: fetal echogenic bowel with one or no detectable CFTR mutation. *J Med Genet.* 2004;41:e70.
9. Ogino S, Flodman P, Wilson RB, Gold B, Grody WW. Risk calculations for cystic fibrosis in neonatal screening by immunoreactive trypsinogen and CFTR mutation tests. *Genet Med.* 2005;7:317–327.
10. Grody WW, Cutting GR, Klinger KW, Richards CS, Watson MS, Desnick RJ. Laboratory standards and guidelines for population-based cystic fibrosis carrier screening. *Genet Med.* 2001;3:149–154.
11. Richards CS, Bradley LA, Amos J, et al. Standards and guidelines for CFTR mutation testing. *Genet Med.* 2002;4:379–391.
12. Watson MS, Cutting GR, Desnick RJ, et al. Cystic fibrosis population carrier screening: 2004 revision of American College of Medical Genetics mutation panel. *Genet Med.* 2004;6:387–391.
13. Wirth B, Herz M, Wetter A, et al. Quantitative analysis of survival motor neuron copies: identification of subtle SMN1 mutations in patients with spinal muscular atrophy, genotype-phenotype correlation, and implications for genetic counseling. *Am J Hum Genet.* 1999;64:1340–1356.
14. Ogino S, Wilson RB, Gold B. New insights on the evolution of the SMN1 and SMN2 region: simulation and meta-analysis for allele and haplotype frequency calculations. *Eur J Hum Genet.* 2004;12:1015–1023.