

## ARTICLE

# EMQN best practice guidelines for the laboratory diagnosis of osteogenesis imperfecta

Fleur S van Dijk<sup>1</sup>, Peter H Byers<sup>2</sup>, Raymond Dalglish<sup>3</sup>, Fransiska Malfait<sup>4</sup>, Alessandra Maugeri<sup>1</sup>, Marianne Rohrbach<sup>5</sup>, Sofie Symoens<sup>4</sup>, Erik A Sidermans<sup>1</sup> and Gerard Pals<sup>\*,1</sup>

Osteogenesis imperfecta (OI) comprises a group of inherited disorders characterized by bone fragility and increased susceptibility to fractures. Historically, the laboratory confirmation of the diagnosis OI rested on cultured dermal fibroblasts to identify decreased or abnormal production of abnormal type I (pro)collagen molecules, measured by gel electrophoresis. With the discovery of *COL1A1* and *COL1A2* gene variants as a cause of OI, sequence analysis of these genes was added to the diagnostic process. Nowadays, OI is known to be genetically heterogeneous. About 90% of individuals with OI are heterozygous for causative variants in the *COL1A1* and *COL1A2* genes. The majority of remaining affected individuals have recessively inherited forms of OI with the causative variants in the more recently discovered genes *CRTAP*, *FKBP10*, *LEPRE1*, *PLOD2*, *PIIB*, *SERPINF1*, *SERPINH1* and *SP7*, or in other yet undiscovered genes. These advances in the molecular genetic diagnosis of OI prompted us to develop new guidelines for molecular testing and reporting of results in which we take into account that testing is also used to 'exclude' OI when there is suspicion of non-accidental injury. Diagnostic flow, methods and reporting scenarios were discussed during an international workshop with 17 clinicians and scientists from 11 countries and converged in these best practice guidelines for the laboratory diagnosis of OI.

*European Journal of Human Genetics* (2012) 20, 11–19; doi:10.1038/ejhg.2011.141; published online 10 August 2011

**Keywords:** EMQN; best practice; osteogenesis imperfecta; type I (pro)collagen; diagnosis; reporting

## INTRODUCTION

### Classification of osteogenesis imperfecta

Osteogenesis imperfecta (OI) comprises a heterogeneous group of disorders characterized by susceptibility to bone fractures with severity that ranges from death in the perinatal period to subtle increase in fracture frequency and in almost all cases presumed or proven defects in type I (pro)collagen biosynthesis.<sup>1</sup> In 1979, David Silience developed a four-type classification, which is still in use for classification according to clinical/radiological features: OI type I (mild OI with bone fragility and blue sclerae), II (perinatal lethal), III (progressive deforming) and IV (normal sclerae and mild deformity).<sup>2,3</sup> Dominant *COL1A1/2* variants appeared to be causative in the majority of OI types. In 2004, OI types V and VI were added to this classification because of specific clinical/radiological and/or histological features, absence of abnormalities of type I (pro)collagen synthesis or structure on gel electrophoresis and absence of causative variants in the *COL1A1/2* genes.<sup>4,5</sup> With the discovery of rare recessive genetic causes of OI (see Table 1), it was proposed to extend the classification with OI types VII and VIII.<sup>6</sup> However, the classification and subdivision into different types of OI is still under discussion<sup>1</sup> because the phenotypic spectrum that results from mutations in some of these genes is almost as broad as that with mutations in type I collagen genes. There is also debate as to whether Bruck syndrome type I<sup>7</sup>

and II,<sup>8</sup> characterized clinically by bone fragility and congenital contractures of the large joints, should be classified as a type of OI.

### Biosynthesis of type I collagen

Most individuals affected with OI are heterozygous for a causative variant in either of the two genes, *COL1A1* or *COL1A2*, which encode the  $\alpha 1(I)$  and  $\alpha 2(I)$  chains of type I procollagen, respectively. Type I collagen is the major structural protein in bone, tendon and ligament. Recently, rare recessive genetic causes of OI have been described and, in all but one, the identified genes encode proteins involved in the biosynthesis of type I procollagen (Figure 1, Tables 1 and 2).<sup>9,10</sup>

### Genes, proteins and causative variants

**Autosomal dominant OI.** The most common form of OI in most populations is OI type I. Cultured dermal fibroblasts from affected individuals produce about half the normal amount of type I procollagen molecules, which have normal structure. OI type I usually results from variants in one *COL1A1* allele (frameshift, nonsense and splice-site alterations) that lead to mRNA instability and haploinsufficiency. In a small subset of individuals with OI type I, substitutions for glycine by small amino acids (cysteine, alanine and serine) near the amino terminal ends of the triple-helical domains of

<sup>1</sup>Department of Clinical Genetics, VU University Medical Centre, Amsterdam, The Netherlands; <sup>2</sup>Department of Pathology and Medicine, University of Washington, Seattle, WA, USA; <sup>3</sup>Department of Genetics, University of Leicester, Leicester, UK; <sup>4</sup>Center for Medical Genetics, Ghent University Hospital, Ghent, Belgium; <sup>5</sup>Division of Metabolism, University Children's Hospital, Zurich, Switzerland

\*Correspondence: Dr G Pals, Department of Clinical Genetics, VU University Medical Centre, PO Box 7057, Amsterdam, 1007 MB, The Netherlands. Tel: +31 20 444 0150; Fax: +31 20 444 0769; E-mail: fs.vandijk2@vumc.nl

This paper was presented at an EMQN Best Practice Meeting, 28–29 June 2010, Amsterdam, The Netherlands. Other participants at the EMQN Best Practice Meeting were Javier Garcia-Planells, Filomena Valentina Gentile, Anne-Sophie Lebre, Shirley McQuaid, Rebecca Pollitt, Agnieszka Rusinska, Thomas Schwarzbraun and Janneke Weiss. Affiliations can be found in Supplementary Appendix Table 2.

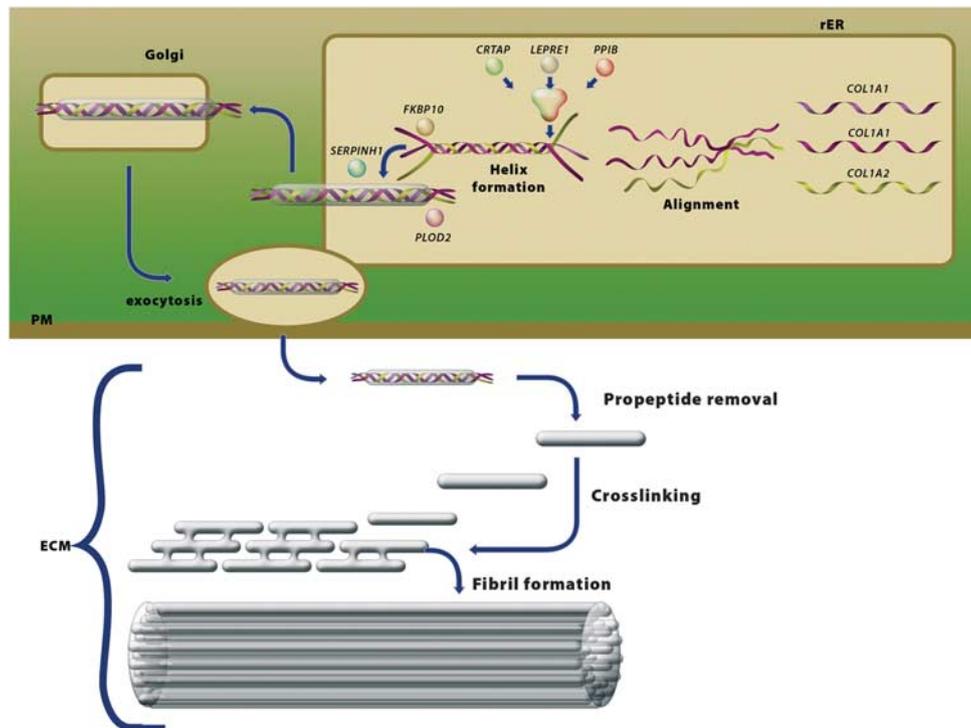
Received 4 February 2011; revised 9 May 2011; accepted 3 June 2011; published online 10 August 2011

**Table 1** Genes in which sequence variants cause OI

Gene	OMIM (Gene)	RefSeqGene genomic reference	RefSeq mRNA reference	RefSeq protein reference	Locus reference genomic (LRG)
<i>COL1A1</i>	120150	NG_007400.1	NM_000088.3	NP_000079.2	LRG_1
<i>COL1A2</i>	120160	NG_007405.1	NM_000089.3	NP_000080.2	LRG_2
<i>CRTAP</i>	605497	NG_008122.1	NM_006371.4	NP_006362.1	LRG_4
<i>FKBP10</i>	607063	NG_015860.1	NM_021939.3	NP_068758.3	LRG_12
<i>LEPRE1</i>	610339	NG_008123.1	NM_022356.3 (transcript variant 1) NM_001146289.1 (transcript variant 2)	NP_071751.3 (isoform 1, P3H1a, long) NP_001139761.1 (isoform 2, P3H1b, short)	LRG_5
<i>PLOD2</i>	601865	NG_009251.1	NM_182943.2 (transcript variant 1) NM_000935.2 (transcript variant 2)	NP_891988.1 (isoform 1, LH2b, long) NP_000926.2 (isoform 2, LH2a, short)	— <sup>a</sup>
<i>PPIB</i>	123841	NG_012979.1	NM_000942.4	NP_000933.1	LRG_10
<i>SERPINF1</i>	172860	NG_028180.1	NM_002615.5	NP_002606.3	— <sup>a</sup>
<i>SERPINH1</i>	600943	NG_012052.1	NM_001235.2	NP_001226.2	— <sup>a</sup>
<i>SP7</i>	606633	NG_023391.1	NM_001173467.1 (transcript variant 1) NM_152860.1 (transcript variant 2)	NP_001166938.1 <sup>b</sup> NP_690599.1 <sup>b</sup>	— <sup>a</sup>

<sup>a</sup>LRG sequences have been developed to provide a stable genomic reference for each gene region annotated with transcripts, proteins and other associated data.<sup>44</sup> For some recently discovered genes in which sequence variation causes OI, no LRG is available yet.

<sup>b</sup>The two transcripts of the *SP7* gene are translated into identical proteins.



**Figure 1** Critical steps in collagen type I biosynthesis and indication of genes known to be involved in OI. Type I collagen is the major structural protein in bone, tendon and ligament. It is first synthesized in the rough endoplasmic reticulum (rER) as type I procollagen, containing C- and N-terminal propeptides. In the rER, the two  $\alpha 1(I)$ -collagen chains encoded by *COL1A1* and the one  $\alpha 2(I)$ -collagen chain encoded by *COL1A2* comprising predominantly Gly-X-Y triplets, align and assemble in the C- to-N direction to form a triple helix. During folding, collagen is modified by, among others, specific enzymes that hydroxylate lysine and proline residues and glycosylate hydroxylysyl residues. This process is called post-translational modification and it stops when triple helix assembly is complete. The CRTAP/P3H1/CyPB complex encoded by the *CRTAP*, *LEPRE1* and *PPIB* genes, is responsible for the 3-hydroxylation of P986 (p.P1164 counting from the methionine that initiates translation) but will most likely also act as a *cis-trans* isomerase and a molecular chaperone. FKBP65 encoded by *FKBP10* also acts as a molecular chaperone for type I procollagen. The protein product of *PLOD2* (procollagen-lysine, 2-oxoglutarate 5-dioxygenase 2) hydroxylates telopeptide lysines in the rER. HSP47 encoded by *SERPINH1* is thought to maintain the stability of the triple helix. After folding, the procollagen molecules are transported through the Golgi apparatus and plasma membrane (PM) into the extracellular matrix (ECM) where cleavage of the N- and C-terminal propeptides occurs and collagen molecules aggregate to form fibrils.

either *COL1A1* or *COL1A2* are found. In contrast, OI types II–IV are usually caused by sequence variants in either *COL1A1* or *COL1A2* that result in substitutions for glycine residues in the uninterrupted

Gly-X-Y triplet repeat of the 1014-residue triple-helical domains encoded by each gene. Less common causative variants include splice-site alterations, variants in the carboxyl-terminal propeptide

**Table 2** Characteristics of recessive OI-related genes

Gene	Protein	Protein function	Results of biochemical analysis in case of causative variants	OI type
<i>CRTAP</i> <sup>18–21,26</sup>	Cartilage-associated protein (CRTAP)	CRTAP forms a complex with the proteins encoded by <i>LEPRE1</i> and <i>PPIB</i> . Apart from 3-prolyl hydroxylation activity, the complex appears to have <i>cis-trans</i> isomerase activity and is thought to act as a molecular chaperone initiating chain recognition and helical folding.	Post-translational over-modification Decreased 3-prolyl hydroxylation of P986 in the pro $\alpha$ 1(I)-collagen chains <sup>a</sup>	II, III, IV
<i>FKBP10</i> <sup>27–29</sup>	FKBP65	FKBP65 has prolyl <i>cis-trans</i> isomerase activity, and can bind collagen	No post-translational over-modification	III/Bruck syndrome (bone fragility with congenital contractures of large joints)
<i>LEPRE1</i> <sup>18–22,23,26</sup>	Prolyl 3-hydroxylase 1 (P3H1)	P3H1 forms a complex with the proteins encoded by <i>CRTAP</i> and <i>PPIB</i> . Apart from 3-prolyl hydroxylation activity, the complex is thought to act as a molecular chaperone and a <i>cis-trans</i> isomerase.	Post-translational over-modification Decreased 3-prolyl hydroxylation of P986 in the pro $\alpha$ 1(I)-collagen chains <sup>a</sup>	II, III
<i>PLOD2</i> <sup>8,31</sup>	Lysyl hydroxylase 2	Telopeptide lysyl hydroxylase-2 modifies lysyl residues in the telopeptides of type I collagen involved in intermolecular cross-link formation	Underhydroxylation of lysine residues in collagen type I telopeptides resulting in aberrant bone collagen type I cross linking	Bruck syndrome
<i>PPIB</i> <sup>18,24–26</sup>	Cyclophilin B	P3H1 forms a complex with the proteins encoded by <i>LEPRE1</i> and <i>CRTAP</i> . Apart from 3-prolyl hydroxylation activity, the complex is thought to act as a molecular chaperone and a <i>cis-trans</i> isomerase	Post-translational over-modification in most cases Decreased 3-prolyl hydroxylation of P986 in the pro $\alpha$ 1(I)-collagen chains in most cases <sup>a</sup>	II, III, IV
<i>SERPINF1</i> <sup>33</sup>	PEDF	PEDF is known mainly for its strong inhibition of angiogenesis. However, expression analyses in bone tissue from wild-type mice and <i>in vitro</i> experiments with murine cell systems support a role for PEDF in bone formation and remodelling	No post-translational over-modification	III
<i>SERPINH1</i> <sup>30</sup>	HSP47	HSP47 monitors the integrity of the triple helix of the type I procollagen at the ER/ <i>cis</i> -Golgi boundary	No post-translational over-modification	III
<i>SP7</i> <sup>32</sup>	Osterix	Osterix is an osteoblast-specific transcription factor which has been shown to be essential for bone formation in mice	No post-translational over-modification	IV

<sup>a</sup>p.Pro1164 counting from the methionine that initiates translation.

coding-domains or insertion/deletion events that lead to in-frame sequence alterations. Most of these variants result in synthesis of abnormal type I procollagen molecules characterized by post-translational over-modification of the triple-helical domain, which results in alterations visible by SDS-polyacrylamide gel electrophoresis. Loss of expression of either gene because of large-scale deletions appears to be rare.<sup>11–15</sup> More than 1000 distinct variants in the *COL1A1* and *COL1A2* genes have been identified that cause OI (R Dalglish: Osteogenesis Imperfecta Variant Database (<https://oi.gene.le.ac.uk>, accessed 3 May 2011)).<sup>16,17</sup>

#### Autosomal recessive OI including Bruck syndrome

In the last 5 years, candidate-gene approaches have identified several loci in which causative variants have been identified which result in the long-sought causes for recessively inherited forms of OI (see Figure 1, Table 2). We classify these genes in four groups—(i) one in which the genes encode proteins that contribute to the initial phase of chain recognition and propagation of molecular folding (*CRTAP*, *LEPRE1* and *PPIB*)<sup>18–26</sup> the second in which the genes encode proteins involved in the final quality control of type I procollagen (*FKBP10*<sup>27–29</sup> and *SERPINH1*<sup>30</sup>), the third, which involves a gene encoding proteins involved in late modification and crosslink formation (*PLOD2*<sup>8,31</sup>) and the fourth group, which involves recently recognized genes

encoding proteins involved in bone cell differentiation (*SP7*<sup>32</sup> and possibly *SERPINF1*<sup>33</sup>).

#### Reasons for referral

**Prenatal.** Prenatal ultrasounds showing abnormalities suggestive of OI (diminished mineralization of skull, platyspondyly or bowing, shortening and/or fractures of long bones) are an important reason for referral for molecular diagnostics of OI.

**Postnatal.** Radiographs at birth suggestive of OI, recurrent and/or unexplained fractures with or without suspicion of non-accidental injury (NAI), primary (idiopathic) low bone mass, preferably with exclusion of secondary causes, a family history of OI and request for confirmation of the clinical diagnosis are all reasons for referral for molecular diagnostics of OI. However, in some cases dentinogenesis imperfecta or even blue sclerae might be the only reason for referral. As such, physicians from many specialties (gynaecologist, paediatrician, orthopaedic surgeon, clinical geneticist, general practitioner, dentist, ear nose throat specialist, endocrinologist, internist and ophthalmologist) might refer a patient for molecular analysis. Depending on the age of presentation, OI can be difficult to distinguish from some other genetic conditions, for example, Ehlers–Danlos syndrome arthrochalasia type (former EDS VIIA and B), isolated

dentinogenesis imperfecta, blue sclerae and corneal fragility, hypophosphatasia, and non-genetic causes of fractures including NAI and idiopathic juvenile osteoporosis.<sup>2,34–36</sup>

**Non-accidental injury.** Many referrals for OI diagnostics occur in the context of suspected NAI in an attempt to exclude a genetic cause of fractures. Fractures resulting from NAI occur in 24 per 10 000 children under 3 years of age whereas the OI prevalence is 1 per 10 000–20 000.<sup>34</sup> The incidence of OI among children evaluated for NAI is 2–5%.<sup>34</sup> Differentiation is aided by an experienced clinician and radiologist familiar with OI<sup>36</sup> as the nature and localization of fractures in OI and NAI can often be distinguished.<sup>37</sup> However, because injuries may involve infants before many of the clinical features of OI are apparent, laboratory diagnostics for OI can certainly be helpful in this differentiation.

## MATERIALS AND METHODS

A group of clinicians and scientists involved in OI diagnostics met on 28–29 June 2010 at a workshop in Amsterdam, the Netherlands, to formulate Best Practice Guidelines for the molecular genetic diagnosis of OI supported by the European Molecular Genetics Quality Network (EMQN).

## RESULTS

Discussions focussed on diagnostic flow, methodologies, interpretation of results and reporting. Consensus guidelines were established.

## DISCUSSION

### The diagnostic flow

The 'traditional' way to establish or confirm the diagnosis of OI is based on studies of collagens synthesized by cultured dermal fibroblasts. Previous studies<sup>35</sup> demonstrated that either quantitative or qualitative alterations can be identified in about 90%<sup>38</sup> of individuals with clinically confirmed OI. Biochemical analysis will separate individuals with quantitative defects (OI type I), from those with qualitative defects (OI types II–IV) because of causative variants in the *COL1A1/2*, *CRTAP*, *LEPRE1*, *PPIB* genes and those with no abnormalities detected. Electrophoretic analysis of type I procollagen may also identify other disorders that can mimic some aspects of OI: EDS kyphoscoliotic type (type VIA), EDS arthrochalasia type (types VIIA and VIIB) and EDS dermatosparaxis type (type VIIC).<sup>11</sup> Also, cultured fibroblasts provides a resource for the analysis of RNA splicing and unclassified variants. However, biochemical analysis will not identify some quantitative defects of type I procollagen, certain causative variants that alter sequences in some coding regions of the *COL1A1/COL1A2* genes and recessive forms of OI (including Bruck syndrome) that result from variants in *FKBP10*, *PLOD2*, *SERPINF1*, *SERPINH1* or *SP7*. Furthermore, prenatal diagnosis by biochemical analysis is only possible in chorionic villus cells and is effective with qualitative alterations<sup>9</sup> but unsuitable for quantitative defects and delays time of diagnosis by 2–4 weeks compared with genetic analysis.

In contrast, direct genomic analysis (sequencing) of the known genes should identify causative variants in >95% of affected individuals in most populations.<sup>39,40</sup> Next generation sequencing has gained importance in the laboratory diagnosis of OI – in part because of shorter time to diagnosis and in part because of the added value of the information as well as reduced costs, which is important for the precise determination of recurrence risk (autosomal dominant versus recessive), prenatal diagnosis and pre-implantation genetic diagnosis.

Given these considerations and current facilities, the consensus of the EMQN Best Practice in OI meeting was to initiate laboratory-based diagnostic studies with direct genomic sequencing of the type I procollagen genes, *COL1A1* and *COL1A2*. The approach to diagnosis is detailed in Figure 2. It was agreed that the 'traditional' approach can still be used in the context in which genomic DNA sequencing is unavailable or the financial barriers are high. Moreover, analysis of proteins and mRNA/cDNA from cultured fibroblasts remain valuable tools for follow-up studies.

### Explanation of diagnostic work flow

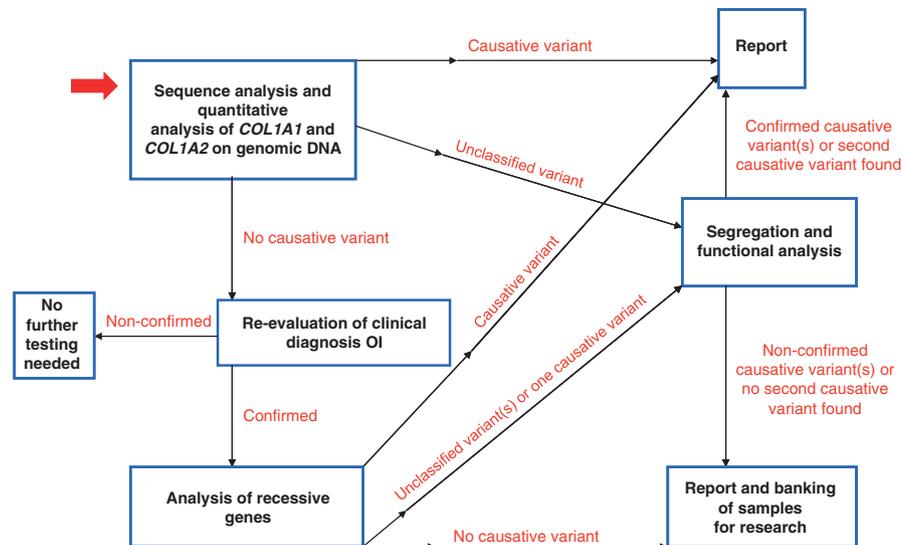
**Sequencing of all type I procollagen gene exons.** Procollagen type I gene sequencing should identify causative variants in 90% of affected individuals,<sup>39,40</sup> provided that the clinical diagnosis of OI is accurate. In some cases a follow-up study is needed to determine whether a variant is causative. This will often include genetic studies in the parents with correlation to the phenotype observed in the parent, or analysis of mRNA splicing and protein-based biochemical studies on cultured dermal fibroblasts.

If no causative variant in the *COL1A1/2* genes is identified by sequence analysis, the next step is to determine if a deletion or duplication of some or all of the coding regions of either gene has occurred. Strategies such as array-based analysis, MLPA or qPCR if properly validated are considered equivalent by the working group in their detection of such alterations. From currently available data in the represented laboratories, the added causative variants expected from this approach should be about 1–2%. In case of OI types II–IV, after identifying the causative variant in the index patient, determination of parental mosaicism by sequence analysis of DNA from blood from each parent will provide data for genetic counselling with respect to recurrence risk and care in subsequent pregnancies.

**Re-review of clinical data and the question of NAI.** When no causative *COL1A1/2* variant is found, clinical priorities should be re-considered. Failure to find a causative variant occurs if there is technical failure, or if no causative variant is present in these genes (because of causative variants in other (un)known genes or because the patient does not have OI). Therefore, referring physicians should be encouraged to provide a completed clinical checklist (Supplementary Table 1) and X-rays of the patient. This clinical/radiological information should be reviewed by a clinical geneticist or a clinician with experience in OI. This can help to determine the likelihood of OI in the particular patient. With clear evidence of OI, further analysis should proceed (Figure 2).

However, if the primary reason for genetic study is to identify children with OI among the larger group suspected for NAI, analysis can reasonably stop after *COL1A1/2* sequencing unless there is a strong suggestion of consanguinity or recessive inheritance. Justification of this decision rests on the following considerations. First, previous studies indicate that fewer than 5% of infants studied for suspicion of NAI are found to have OI by biochemical or DNA-based studies.<sup>34</sup> Second, DNA-based analysis will identify a causative variant in >90% of all individuals with OI<sup>39,40</sup> so that the remaining risk that an infant has OI, will be about 0.5%. Third, at present all infants with recessive OI have obvious radiographic abnormalities fitting a diagnosis of OI and not of NAI.

**Identification and characterization of causative variants in autosomal recessive OI.** Variants in the genes causing recessive OI are estimated to account for about 5 or 6% of individuals with OI. This represents the pooled estimates from all the laboratories represented at the workshop. Although laboratory context and clinical criteria may in



**Figure 2** Preferred diagnostic flow in OI. The approach to diagnosis is designed to maximize the likelihood that causative variants will be identified in all affected individuals or assign those without causative variants to research pools. This flow assumes that the clinical diagnosis of OI is well established according to the traditional diagnostic criteria. With clear evidence of OI from radiological and clinical examination, further analysis should proceed according to the proposed strategy. Functional analysis consists of analysis of proteins and mRNA/cDNA from cultured fibroblasts and also includes *COL1A1* null allele testing in certain selected cases.

some cases require otherwise, the preferred strategy is to analyse all genes at the same time to minimize turnaround time. If no, or only one, causative variant is identified by sequence analysis, the next step is to determine by array-based analysis, MLPA or qPCR if a deletion or duplication of some or all of the coding regions of either allele has occurred. Once recessive causative variants are identified, parental confirmation of carrier status should be completed, first to confirm that the causative variants in the index case are inherited in *trans* and, second, to exclude the possibility of uniparental disomy in the case of homozygosity in the infant.

*Analysis of proteins and mRNA/cDNA from cultured fibroblasts (functional analysis).* By following the guidelines to this point, virtually all causative variants in type I procollagen genes and in the recessive OI-related genes will have been identified. However, in some cases analysis of proteins and mRNA/cDNA from cultured fibroblasts can have an additive value. First of all, mRNA/cDNA analysis provides a tool for studying the effect of unclassified variants suspected to alter splicing. Moreover, this approach can unravel the effect of deep intronic variants that alter the mRNA sequences usually by conversion of cryptic exons to active exons and which would not be detected by gDNA sequencing. In these cases, biochemical analysis of type I procollagen chains can detect quantitative defects if the new exon results in introduction of a premature stop codon and nonsense-mediated mRNA decay, or qualitative defects if the mRNA is stable and translated into an abnormal protein. When OI type I is suspected, but genomic DNA sequencing is unavailable or too costly, *COL1A1* null-allele<sup>14</sup> testing on cDNA can be used to confirm the diagnosis. However, this approach relies on the availability of informative markers and will not lead to the identification of the disease-causing variant.

*Analysis of remaining samples.* Some diagnostic laboratories maintain research arms and the identification of the additional OI-related genes has profited from the repositories of cells and DNA samples from individuals with OI because of unknown genetic causes. This

practice should continue and cells and DNA samples should be banked for future analysis. If the diagnostic laboratory does not maintain such a resource, banking with ones that do, should be considered.

### Laboratory diagnosis of OI

*Molecular analysis for OI.* Molecular analysis for OI includes sequencing or mutation scanning and deletion/duplication testing for *COL1A1* and *COL1A2*, followed by sequencing of recessive OI-related genes (see Table 1). Preferred material for testing is genomic DNA from blood, but other sources can be used as well (see Tables 3a and b). For general issues regarding sample labelling and identification, sequence analysis and interpretation see 'Practice guidelines for Sanger Sequencing Analysis and Interpretation' (<http://www.cmg.org/BPGs/pdfs%20current%20bpgs/Sequencing2.pdf>).

### gDNA and cDNA sequencing

Current practices in sequencing of the type I collagen genes differ among laboratories in that some use exon-by-exon amplification followed by sequence determination (Supplementary Appendix Tables 3A and B) and others use a multi-exon substrate for analysis. If properly validated, both will provide the same mutation capture probability and their use should be determined by local preferences. To identify splice variants, primer sequences for PCR used for direct sequencing should be designed far enough from intron–exon boundaries to allow reading of the branch sites, consensus splice donor and acceptor sites or UTR sequences of acceptable quality. With regard to sequence analysis of cDNA, large fragments should be analysed to allow detection of multi-exon deletions, because these could be missed when primers are located close to the exon junctions at multiple sites. The analysis of cDNA from fibroblasts is warranted when an unknown variant is found that might affect splicing. In addition, cDNA sequence analysis may detect null-alleles without a known cause (eg, an unknown variant in the promoter region) if care is taken to analyse regions in which coding sequence heterozygosity is known.

**Table 3a Postnatal diagnostics for OI**

Source	Product	Methods	Comments
Blood (preferred)	gDNA	Direct sequencing Mutation scanning (HRM/QPCR) Deletion/duplication testing (MLPA)	
Blood spots	gDNA	Direct sequencing	Very low yield
Saliva or buccal swabs	gDNA	Direct sequencing Mutation scanning (HRM/QPCR) Deletion/duplication testing (MLPA)	High yield of DNA possible
Fibroblasts	gDNA	Direct sequencing Mutation scanning (HRM/QPCR) Deletion/duplication testing (MLPA) <sup>a</sup>	
	mRNA/cDNA	Sequencing for splice-site errors	
	Protein	(Pro)collagen type I electrophoresis	

Abbreviation: HRM, high-resolution melting.

<sup>a</sup>Performing MLPA with DNA extracted from fibroblasts leads to unclear results in certain cases.

**Table 3b Prenatal and preimplantation genetic diagnosis for OI**

Source	Product	Methods
CVS	gDNA	Direct sequencing Mutation scanning (HRM/QPCR) Deletion/duplication testing (MLPA)
	Protein	(Pro)collagen type I electrophoresis
Amniocytes	gDNA	Direct sequencing Mutation scanning (HRM/QPCR) Deletion/duplication testing (MLPA)
Blastocysts	gDNA	Sequencing of known familial mutations

### Quantitative analysis

Quantitative analysis by MLPA (kits for *COL1A1* and *COL1A2* have been designed by MRC-Holland, available at <http://www.mlpa.com>) and qPCR is particularly important in cases without a detected molecular cause in which only genomic DNA is available and without heterozygous polymorphic variants detected by sequencing. Recently, complete *COL1A1* allele deletions have been reported to cause OI type I.<sup>13</sup> For reliable quantitative testing (MLPA/qPCR) patient and control materials from the same source (eg, blood or fibroblasts) should be used for comparison and the same DNA purification protocol should be used for all samples within a test series.

**Protein analysis.** Protein analysis of type I (pro)collagen (Supplementary Appendix Figures 3A–D) is used to detect quantitative and qualitative changes. The method used for this analysis is based on *in vitro* labelling of collagen in cultured fibroblasts, followed by electrophoresis on 2 M urea SDS-PAGE gels (the 2 M urea enhances chain separation) and autoradiography.<sup>41</sup> The analysis of procollagens is achieved by omitting the pepsin digestion step from the collagen protocol. Procollagen electrophoresis enhances detection sensitivity of quantitative type I (pro)collagen defects (OI type I), visualizes defects in type I (pro)collagen located in the N-terminal region of the type I collagen triple helix more clearly and distinguishes between EDS VIIC and EDS VIIA and B (if samples are analysed after labelling in the

presence and absence of dextran sulphate to enhance proteolytic processing).<sup>42</sup>

**Prenatal diagnosis.** Prenatal diagnosis is possible in case of identification of known disease-causing variant(s) both on genomic DNA extracted from chorionic villus sample (CVS) cells and amniocytes. ‘Testing for Maternal Cell Contamination (MCC) in Prenatal Samples for Molecular Studies’ ([http://www.cmgs.org/BPGs/pdfs%20current%20bpgs/MCC\\_08.pdf](http://www.cmgs.org/BPGs/pdfs%20current%20bpgs/MCC_08.pdf)) should be applied. In those rare instances when post-translational over-modification of (pro)collagen type I is visible in, for example, cells from an affected sibling but no disease-causing variant(s) have been detected in the genes involved in OI, electrophoresis of type I (pro)collagen from cultured CVS cells can be used for prenatal diagnosis. Amniocytes are good sources of DNA but they do not make type I procollagen in sufficient abundance to permit prenatal diagnosis.

### Interpretation of performed diagnostics

For the interpretation of an observed sequence variant it is essential to establish the causal role of the variant in the pathogenesis of the disease. Textboxes 1 and 2 include descriptions of variants in the genes involved in OI that are likely to be pathogenic or that should be considered as unclassified variants.

### Reporting

General information on requirements for variant reporting can be found in the OECD Guidelines for Quality Assurance in Molecular Genetic Testing (<http://www.OECD.org/dataoecd/43/6/38839788.pdf>) and in the guidelines issued by the Swiss Medical Genetics Society ([http://sgmg.ch/user\\_files/images/SGMG\\_Reporting\\_Guidelines.pdf](http://sgmg.ch/user_files/images/SGMG_Reporting_Guidelines.pdf)).

### Reporting scenarios

**Finding a causative variant in the *COL1A1*/*COL1A2* genes in an affected index case.** The report should state that a causative variant has been detected and that this confirms the clinical diagnosis. The report should include a description of the reason why a particular variant is considered causative. Determination of the parental origin of the detected variant(s) should be advised. Testing of relatives at risk should be offered in conjunction with appropriate counselling.

In the case of severe/lethal OI because of a causative variant in *COL1A1* or *COL1A2*, the measured recurrence risk is 2% after the birth of one affected child. The recurrence risk is increased after the birth of two affected individuals, presumably because the proportion of germ cells that carry the mutation is higher.<sup>43</sup> This information may be mentioned in the results letter. If parental mosaicism can be demonstrated in the father, the referring clinician could be encouraged to request a sperm sample from the father to clarify the risk.

When the referring physician is not a clinical geneticist, it is recommended that the patient be referred to a clinical genetics centre for counselling. Referral to OI centres for therapy can be proposed if available.

**Not finding a causative variant in *COL1A1* and *COL1A2* genes in an affected index case.** If no causative variant is detected in the *COL1A1* and *COL1A2* genes after gDNA analysis and screening for large gene deletions (eg, MLPA), an *interim report* can be made with the recommendation for review of the original patient diagnosis (detailed clinical information, eg, with use of the clinical checklist, radiographs, contact with the referring physician (see Supplementary Appendix Table 1)). This review of the clinical diagnosis should involve a clinician with expertise in diagnosis and classification of OI. The

outcome of this review might indicate the desirability for testing of additional OI genes and cDNA and protein analysis of type I (pro)collagen. If applicable, a new request for additional testing should be sent by the referring physician to the laboratory. Laboratories that do not offer analysis of the recessive genes should suggest further analysis in another laboratory.

In the case of doubtful pathogenicity of the identified variant, the 'Practice guidelines for the Interpretation and Reporting of Unclassified Variants (UV's) in Clinical Molecular Genetics' (<http://www.cmgs.org/BPGs/pdfs%20current%20bpgs/UV%20GUIDELINES%20ratified.pdf>) should be applied. If the genetic testing procedure has identified an unclassified variant as the only sequence change, this should be reported as such. However, the report should clearly state that the clinical significance of the variant is unknown and that its identification does not provide an explanation for the clinical phenotype of the patient.

*Finding causative variants in the CRTAP, LEPRE1, PPIB, FKBP10, SERPINF1, SERPINH1, PLOD2, SP7 genes in an affected index case.* The final report should state that causative variants have been detected. It is necessary to investigate the carrier status of the parents and siblings in order to determine whether a variant is causative. Testing of relatives at risk may be offered in conjunction with appropriate counselling. Such testing might be useful for prenatal diagnosis.

*Finding one causative variant in the CRTAP, LEPRE1, PPIB, FKBP10, SERPINF1, SERPINH1, PLOD2, SP7 genes in an affected index case.* When one causative recessive variant has been detected, it cannot be excluded that a second causative variant has been missed, for example when it concerns a deep intronic variant. Dependent on the laboratory, a second interim report can be made indicating the desirability for cDNA and protein analysis of type I (pro)collagen in the case of one causative variant in the CRTAP, LEPRE1 or PPIB genes. When these additional tests are not informative or cannot be performed in the laboratory, the final report should state that it is not possible to confirm the clinical diagnosis OI by biochemical and/or molecular testing.

*Not finding causative variants in the CRTAP, LEPRE1, PPIB, FKBP10, SERPINF1, SERPINH1, PLOD2, SP7 genes in an affected index case.* In the case of doubtful pathogenicity of the identified variants, the same procedure as described above for the COL1A1 and COL1A2 genes should be followed. In the case where no causative variants in the genes involved in recessive OI have been found, the final report should state that it is not possible to confirm the clinical diagnosis OI by biochemical and/or molecular testing.

*Not finding causative variant(s) in the genes known for autosomal dominant and autosomal recessive OI: type I (pro)collagen electrophoresis.* Reports of these studies are necessarily more descriptive than the DNA-based results. They should list the types of genetic alterations that could not be identified. When low production or post-translational over-modification of type I (pro)collagen is observed, the diagnosis of OI can be confirmed and in the case of over-modification of type I (pro)collagen, prenatal diagnosis on chorionic villus cells is possible. In reports concerning cells and samples with no molecular and/or biochemical diagnosis of OI, it should be stated that it is not possible to confirm the clinical diagnosis OI.

#### Prenatal or preimplantation genetic diagnosis

Prenatal or preimplantation genetic diagnosis with the intention of terminating a pregnancy or not selecting embryos carrying the

causative variant(s) is possible in the case of identification of known disease-causing variant(s). This service should only be offered in a clinical genetics service and must be accompanied by appropriate genetic counselling. Requests for prenatal or preimplantation diagnosis should always be referred and announced in advance to a clinical genetics service. In addition, the original result letter of the laboratory that identified the causative variant should accompany the request.

#### CONCLUSIONS

Best practice guidelines have been established for the molecular genetic diagnosis of OI. The most noteworthy issue is that molecular analysis of the COL1A1/2 genes is recommended as the starting point in the diagnostic flow as opposed to protein analysis. It is to be expected that molecular genetic testing in OI will only gain in importance as new OI-genes are discovered and with the development of new technologies. However, in certain selected cases protein analysis will remain important.

#### CONFLICT OF INTEREST

The authors declare no conflict of interest.

#### ACKNOWLEDGEMENTS

Funding for the Best Practice Meeting was provided by the European Molecular Genetics Quality Network (<http://www.emqn.org>).

- van Dijk FS, Pals G, van Rijn RR, Nikkels PGJ, Cobben JM: Classification of osteogenesis imperfecta revisited. *Eur J Med Genet* 2010; **53**: 1–5.
- Byers PH, Krakow D, Nunes ME, Pepin M: Genetic evaluation of suspected osteogenesis imperfecta (OI). *Genet Med* 2006; **8**: 383–388.
- Sillence DO, Senn A, Danks DM: Genetic heterogeneity in osteogenesis imperfecta. *J Med Genet* 1979; **16**: 101–116.
- Glorieux FH, Rauch F, Plotkin H *et al*: Type V osteogenesis imperfecta: a new form of brittle bone disease. *J Bone Miner Res* 2000; **15**: 1650–1658.
- Glorieux FH, Ward LM, Rauch F, Lalic L, Roughly PJ, Travers R: Osteogenesis imperfecta type VI: a form of brittle bone disease with mineralization defect. *J Bone Miner Res* 2002; **17**: 30–37.
- Rauch F, Glorieux FH: Osteogenesis imperfecta. *Lancet* 2004; **363**: 1377–1385.
- Breslau-Siderius EJ, Engelbert RH, Pals G, van der Sluijs JA: Bruck syndrome: a rare combination of bone fragility and multiple congenital joint contractures. *J Pediatr Orthop B* 1998; **7**: 35–38.
- Ha-Vinh R, Alanay Y, Bank RA *et al*: Phenotypic and molecular characterization of Bruck syndrome (osteogenesis imperfecta with contractures of the large joints) caused by recessive mutation in PLOD2. *Am J Med Genet* 2004; **131A**: 115–120.
- Engel J, Prockop DJ: The zipper-like folding of collagen triple helices and the effects of mutations that might disrupt the zipper. *Annu Rev Biophys Chem* 1991; **20**: 137–152.
- Byers PH, Wallis GA, Willing MC: Osteogenesis imperfecta: translation of mutation to phenotype. *J Med Genet* 1991; **28**: 433–442.
- Steiner RD, Pepin MG, Byers PH: Osteogenesis imperfecta. Available at <http://www.ncbi.nlm.nih.gov/bookshelf/br.fcgi?book=gene&part=oi>. Accessed 3 May 2011.
- Willing MC, Pruchno CJ, Atkinson M, Byers PH: Osteogenesis imperfecta type I is commonly due to a COL1A1 null allele of type I collagen. *Am J Hum Genet* 1992; **51**: 508–515.
- Willing MC, Deschenes SP, Scott DA *et al*: Osteogenesis imperfecta type I: molecular heterogeneity for COL1A1 null alleles of type I collagen. *Am J Hum Genet* 1994; **55**: 638–647.
- Nuytinck L, Sayli BS, Karen W, De Paeppe A: Prenatal diagnosis of osteogenesis imperfecta type I by COL1A1 null-allele testing. *Prenat Diagn* 1999; **19**: 873–875.
- van Dijk FS, Huizer M, Kariminejad A *et al*: Complete COL1A1 allele deletions in osteogenesis imperfecta. *Genet Med* 2010; **12**: 736–741.
- Dalgleish R: The human type I collagen mutation database. *Nucleic Acids Res* 1997; **25**: 181–187.
- Dalgleish R: The human collagen mutation database 1998. *Nucleic Acids Res* 1998; **26**: 253–255.
- Ishikawa Y, Wirz J, Vranka JA, Nagata K, Bächinger HP: Biochemical characterization of the prolyl 3-hydroxylase 1-cartilage-associated protein-cyclophilin B complex. *J Biol Chem* 2009; **284**: 17641–17647.
- Barnes AM, Chang W, Morello R *et al*: Deficiency of cartilage-associated protein in recessive lethal osteogenesis imperfecta. *N Engl J Med* 2006; **355**: 2757–2764.
- Morello R, Bertin TK, Chen Y *et al*: CRTAP is required for prolyl 3-hydroxylation and mutations cause recessive osteogenesis imperfecta. *Cell* 2006; **127**: 291–304.

- 21 Marini JC, Cabral WA, Barnes AM: Null mutations in *LEPRE1* and *CRTAP* cause severe recessive osteogenesis imperfecta. *Cell Tissue Res* 2010; **339**: 59–70.
- 22 Cabral WA, Chang W, Barnes AM *et al*: Prolyl 3-hydroxylase 1 deficiency causes a recessive metabolic bone disorder resembling lethal/severe osteogenesis imperfecta. *Nat Genet* 2007; **39**: 359–365.
- 23 Willaert A, Malfait F, Symoens S *et al*: Recessive osteogenesis imperfecta caused by *LEPRE1* mutations: clinical documentation and identification of the splice form responsible for prolyl 3-hydroxylation. *J Med Genet* 2009; **46**: 233–241.
- 24 van Dijk FS, Nesbitt IM, Zwikstra EH *et al*: *PPIB* mutations cause severe osteogenesis imperfecta. *Am J Hum Genet* 2009; **85**: 521–527.
- 25 Barnes AM, Carter EM, Cabral WA *et al*: Lack of cyclophilin B in osteogenesis imperfecta with normal collagen folding. *N Engl J Med* 2010; **362**: 521–528.
- 26 Pyott SM, Schwarze U, Christiansen HE *et al*: Mutations in *PPIB* (cyclophilin B) delay type I procollagen chain association and result in perinatal lethal to moderate osteogenesis imperfecta phenotypes. *Hum Mol Genet* 2011; **20**: 1595–1609.
- 27 Alanay Y, Avaygan H, Camacho N *et al*: Mutations in the gene encoding RER protein FKBP65 cause autosomal-recessive osteogenesis imperfecta. *Am J Hum Genet* 2010; **86**: 551–559.
- 28 Shaheen R, Al-Owain M, Sakati N, Alzayed ZS, Alkuraya FS: FKBP10 and Bruck syndrome. Phenotypic heterogeneity or call for reclassification? *Am J Hum Genet* 2010; **87**: 306–307.
- 29 Kelley BP, Malfait F, Bonafe L *et al*: Mutations in *FKBP10* cause recessive osteogenesis imperfecta and Bruck syndrome. *J Bone Miner Res* 2011; **26**: 666–672.
- 30 Christiansen HE, Schwarze U, Pyott SM *et al*: Homozygosity for a missense mutation in *SERPINH1*, which encodes the collagen chaperone protein HSP47, results in severe recessive osteogenesis imperfecta. *Am J Hum Genet* 2010; **86**: 389–398.
- 31 van der Slot AJ, Zuurmond AM, Bardoel AFJ *et al*: Identification of *PLOD2* as telopeptide lysyl hydroxylase, an important enzyme in fibrosis. *J Biol Chem* 2003; **278**: 40967–40972.
- 32 Lapunzina P, Aglan M, Temtamy S *et al*: Identification of a frameshift mutation in *Osterix* in a patient with recessive osteogenesis imperfecta. *Am J Hum Genet* 2010; **87**: 110–114.
- 33 Becker J, Semler O, Gilissen C *et al*: Exome sequencing identifies truncating mutations in human *SERPINF1* in autosomal-recessive osteogenesis imperfecta. *Am J Hum Genet* 2011; **88**: 362–371.
- 34 Marlowe A, Pepin MG, Byers PH: Testing for osteogenesis imperfecta in cases of suspected non-accidental injury. *J Med Gen* 2002; **39**: 382–386.
- 35 Ablin DS, Greenspan A, Reinhart M, Grix A: Differentiation of child abuse from osteogenesis imperfecta. *Am J Roentgenol* 1990; **154**: 1035–1046.
- 36 Steiner RD, Pepin M, Byers PH: Studies of collagen synthesis and structure in the differentiation of child abuse from osteogenesis imperfecta. *J Pediatr* 1996; **128**: 542–547.
- 37 Bilo RAC, Robben SGF, van Rijn RR: *Forensic Aspects of Paediatric Fractures; Differentiating Accidental Trauma from Child Abuse*. 1st edn. Springer: Heidelberg, Dordrecht, London, New York, 2010.
- 38 Wenstrup RJ, Willing MC, Starman BJ, Byers PH: Distinct biochemical phenotypes predict clinical severity in nonlethal variants of osteogenesis imperfecta. *Am J Hum Genet* 1990; **46**: 975–982.
- 39 Körkkö J, Ala-Kokko L, De Paepe A, Nuytinck L, Earley J, Prockop DJ: Analysis of the *COL1A1* and *COL1A2* genes by PCR amplification and scanning by conformation-sensitive gel electrophoresis identifies only *COL1A1* mutations in 15 patients with osteogenesis imperfecta type I: identification of common sequences of null-allele mutations. *Am J Hum Genet* 1998; **62**: 98–110.
- 40 Sykes B, Ogilvie D, Wordsworth P *et al*: Consistent linkage of dominantly inherited osteogenesis imperfecta to the type I collagen loci: *COL1A1* and *COL1A2*. *Am J Hum Genet* 1990; **46**: 293–307.
- 41 Steinmann B, Rao VH, Vogel A, Bruckner P, Gitzelmann R, Byers PH: Cysteine in the triple-helical domain of one allelic product of the  $\alpha 1(I)$  gene of type I collagen produces a lethal form of osteogenesis imperfecta. *J Biol Chem* 1984; **259**: 11129–11138.
- 42 Bateman JF, Golub SB: Assessment of procollagen processing defects by fibroblasts cultured in the presence of dextran sulfate. *Biochem J* 1990; **267**: 573–577.
- 43 Pyott SM, Schwarze U, Christiansen HE *et al*: Recurrence of perinatal lethal osteogenesis imperfecta in sibships: parsing the risk between parental mosaicism for dominant mutations and autosomal recessive inheritance. *Genet Med* 2011; **13**: 125–130.
- 44 Dagleish R, Fliceck P, Cunningham F *et al*: Locus Reference Genomic sequences: an improved basis for describing human DNA variants. *Genome Med* 2010; **2**: 24.



This work is licensed under the Creative Commons Attribution-NonCommercial-No Derivative Works 3.0 Unported Licence. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/3.0/>

Supplementary Information accompanies the paper on European Journal of Human Genetics website (<http://www.nature.com/ejhg>)

## APPENDIX

### Textbox 1

Interpretation of variants in the *COL1A1* and *COL1A2* genes.

#### *COL1A1* and *COL1A2* genes

1. Causative variants – considerations specific for the fibrillar collagen genes:

*The following variants are likely to have pathological consequences for the protein function:*

- a. Substitution of the glycine residue of every Gly-X-Y triplet repeat within the triple-helical region of the  $\alpha 1(I)$ - and  $\alpha 2(I)$ -collagen chains.
- b. Sequence variants in consensus acceptor and donor splice sites (positions +1, +2, –1, –2) will lead to splicing errors of the mRNA, but the outcome cannot be predicted without cDNA analysis, with the exception of –1G>A acceptor-site splicing variants in the triple-helical region of the  $\alpha 1(I)$  collagen chains which result in a frameshift, nonsense-mediated mRNA decay and a null-allele. –1G>A splicing variants of exon 6 of the *COL1A1* gene, results in skipping of exon 6, whereas for the *COL1A2* gene this results in use of a cryptic splice site in the exon and an in-frame deletion of 15 bases. Both are associated with EDS arthrochalasia type.
- c. Nonsense mutations or deletions and insertions that result in interruption of the reading frame, nonsense-mediated mRNA decay and a null-allele.
- d. In-frame deletions and insertions that lengthen or shorten the  $\alpha$ -chains.
- e. Other variants\* with experimental evidence (published or own data) of their impairment of the protein's function.

\*Variants also should be checked against existing entries in the Osteogenesis Imperfecta Variant Database<sup>16,17</sup> (<https://oi.gene.le.ac.uk>), which can be used for corroboration. However, it should be recognized that the data in the database are not warranted to be accurate or fit for any particular purpose.

2. Unclassified variants (UV)

- a. All other sequence variants must be considered as 'unclassified' until segregation within the family has been investigated and/or functional evidence becomes available.
- b. It is strongly recommended that intronic nucleotide substitutions other than splice-site mutations affecting the –1, –2 or +1, +2 intronic nucleotide positions be examined by cDNA-analysis to determine their splicing outcome.
- c. It is strongly recommended that UVs be further investigated by *in silico* analysis (eg, PolyPhen-2, SIFT, Human Splicing Finder software, and so on). Of note: the new version of Polyphen classifies almost all glycine substitutions as benign or tolerated.

## Textbox 2

Interpretation of variants in the *CRTAP*, *FKBP10*, *LEPRE1*, *PLOD2*, *PIIB*, *SERPINF1*, *SERPINH1* and *SP7* genes.

---

*CRTAP*, *FKBP10*, *LEPRE1*, *PLOD2*, *PIIB*, *SERPINF1*, *SERPINH1* and *SP7*

1. Causative variants

*The following variants are likely to have pathological consequences for the protein function:*

- a. Sequence variants in consensus acceptor and donor splice sites that will lead to splicing errors of the mRNA.
- b. Nonsense mutations or deletions and insertions that result in interruption of the reading frame, nonsense-mediated mRNA decay and a null-allele.
- c. Most in frame deletions and insertions that lengthen or shorten the protein.
- d. See 1e in Textbox 1

2. Unclassified variants

See Textbox 1.

---