Frequently Asked Questions from Patients and Contributors

Table of contents

1. What benefits will I receive from contributing my images?

1.1. Will my contribution of images serve the common good?

Yes.

Your images will expand the pool of available information to help medical imaging research, software development and education. This information will be available to everyone in the world via the Internet, without charge or restriction, serving as a global, accessible, community resource.

It is hoped that the long term benefit of the contributions from you and others like you will be the improved quality and available of medical imaging technology for the detection, diagnosis and treatment of disease.

1.2. Will I be paid for my contribution of images?

No.

The PCIR is founded on the assumption that contributors will see long-term value in contributing to the welfare of others. Your generosity is greatly appreciated.

1.3. Will I be recognized publicly for my contribution of images?

No.

Making your identity known as a contributor might put your privacy at risk.

1.4. Will you tell me if you find anything on my images?

No.

The PCIR only de-identifies and distributes your images and does not provide any interpretation of the images. Also, as described in the introductory section here, to protect your privacy, there is no provision for anyone who may download your de-identified images to communicate with you, nor for the PCIR to forward any such communication.

2. What privacy concerns exist?

2.1. What risks to privacy do medical images pose?

Firstly, medical images usually contain information such as your name. Though it may not be visible in the image pixels themselves, the image formats contain a "header" that contains lots of information about when the exam was performed, where, by whom and upon whom.

If this information was not removed before the images are shared with the public, then anyone who cared to look would know that these images belong to you, the contributor.

Secondly, it may be apparent from the appearance of an image, or from the information accompanying an image, that you have a certain disease or have undergone a certain treatment, or are in a certain risk category.

The combination of your identity with the disease or risk status thus creates a potential risk.

2.2. Why does it matter if anyone recognizes the images are of me?

First of all, most people simply like to keep their medical affairs private.

More importantly though, making people aware that you have a particular disease or risk factor may have a negative impact, socially or economically. For example, your family or friends may become unduly concerned, or an employer or insurer could discriminate against you.

For such reasons, it is important to de-identify the images before they are released publicly.

2.3. How does PCIR protect my privacy?

PCIR considers your privacy our number one concern.

For this reason, when you contribute, the images and other files are packaged and encrypted with a secure algorithm, before they leave your computer. The files are then thoroughly "de-identified" by PCIR before being made publicly available.

PCIR does not sell, share or otherwise make available information containing your identity to any third party.

2.4. What information is removed during de-identification?

The PCIR de-identification process involves such things as:

- Removing your name and your identifying information like medical record numbers, addresses, etc.
- Removing your hospital's and doctors' names, addresses, etc.
- Removing any identifying information "burned in" to the image pixels themselves.
- Removing some other types of information that could potential be used to identify you and which have no relevance for research, such as your date of birth, and the dates of the examination.
- Removing any identifiers of specific acquisition devices, such as serial numbers.
- Removing any unique identifiers of images, series and studies.
- Removing any extra information in other text values that is not clearly safe to retain.

- Removing any information of unknown safety such as in manufacturers' private attributes.

2.5. What information is not removed during de-identification?

Some types of information are potentially "sensitive" but need to be retained in order to make the images useful for some research purposes. These include:

- Your physical characteristics at the time of examination, such as sex, age, and height and weight.
- The contents of the images themselves, which in some cases might theoretically be reconstructed to create a likeness of, for example, your face.

2.6. Can I perform de-identification myself before contributing?

Theoretically it is possible for you to de-identify your own images, if they are in the right format, and if you have access to the right software and have the knowledge to use it. PCIR does not currently supply such tools or implement this capability in the upload applet, since we believe it is too burdensome to expect most contributors to do this. For some file formats, such as DICOM, there are tools available in the Internet, but many are of poor quality and actual damage the DICOM header and make it unusable.

Even if you do perform your own de-identification, PCIR will still go through the process of completely re-de-identifying the images and information again anyway.

Furthermore, we prefer to do the de-identification at our end, because we first put the information through processes that allow us to match your contribution with prior or future contributions that you may have made or may make, and to match image contributions with other information such as the scanned paper documents that you submit.

Similarly, for scanned paper documents, we prefer that you do not remove your identifying information prior to scanning, since we use that information to confirm that the documents and the images that you have contributed are of the same patient. PCIR will remove any such information prior to distribution.

2.7. Does contributing create a risk of identity theft?

Identity theft depends on the perpetrator using multiple pieces of information about you, perhaps obtained from multiple sources, to masquerade as you to obtain services.

The de-identification process removes personally identifiable and sensitive information such that there should be no such risk. For example, your name, date of birth, social security number and so on would all be removed.

2.8. Can somebody already in possession of my images recognize them?

Yes. Since the de-identification process does not change the image pixel data itself (except in the unusual case when burned in identification needs to be cut out), it is possible for somebody who already has your images to match the de-identified images that are publicly released, if they were sufficiently determined.

We suggest that this does not matter, since anyone who already has access to your images already has more information from them than may be gleaned from the de-identified version. If you are not comfortable with this, then please do not contribute.

2.9. What if a privacy breach occurs, and identifying information is found in publicly accessible images?

In the unlikely event that this occurs, you should immediately notify the PCIR of the privacy breach by sending an email to privacy@pcir.org.

The offending material will then be removed from further distribution via the PCIR site.

2.10. Is anything saved on my local computer during the contribution process?

The uploading tool creates temporary files on your local hard drive whilst it is compressing and encrypting the images read from the CD and any scanned document files. These temporary files may contain your identifying information. These files are automatically deleted when the tool finishes. They will usually also be deleted even if the tool fails for some reason.

Despite this, we strongly recommend that you use a computer that is under your personal control rather than a public machine (for example in a public library) or a computer at work (which may be accessible to IT staff).

Additional measures to help remove any residual identifying information are described in the troubleshooting FAQ section on scrubbing your computer.

Some additional information is also saved between sessions, in order to be able to recognize that you are the same person, for example if you upload some files today and more tomorrow, or next time you have an imaging examination. This also avoids repeatedly asking you to agree to the same terms and conditions each time you visit the contribution web page within a short period of time. Your true identity is not stored, only the fact that you have contributed before.

There is a troubleshooting FAQ section on where session information is stored if you are concerned about removing this.

3. Secondary Use

3.1. What is "Secondary Use"?

This term has recently come into use to describe the secondary uses to which those who **already have access** to a patient's data for some legitimate and authorized purpose related to provision of healthcare services, may re-use it, or pass it on to other re-use, for some other purpose.

Such purposes include outcomes analysis, research, quality assurance, and public health as well as purely commercial purposes such as selling data for marketing or other business purposes.

Secondary use may or may not involve de-identification of your data. Indeed, some secondary use purposes specifically require your identity, and some people constrain the use of the term to data including your identifiable information.

For more information, you may wish to read some of the externals links on secondary use.

3.2. Should I be concerned about "Secondary Use", and if so why?

Yes.

Secondary Use is not inherently bad, quite the contrary, but you need to be extremely concerned about the type of secondary use. Essentially, many of those interested in gaining access to your identifiable healthcare information may not have your best interests at heart.

Some secondary use purposes are clearly both in your interest and in the public good, such as research and outcomes analysis that are based on de-identified data, or quality assurance that seeks to provide better diagnosis or treatment; presuming of course that the appropriate safeguards to your privacy are in place.

Some purely commercial uses, however, are neither in your interest nor the public good, such as those that seek to deny you access to healthcare or increase its cost based on your risk or deny you employment. Nor, quite likely, do you desire to be inundated with marketing and sales information based on targeting derived from your healthcare information.

3.3. Is contributing to the PCIR a form of "Secondary Use"?

No.

Though the purpose is to "re-use" your images, the PCIR receives contributions directly from you, the subject of of the images and information, and you give express permission for it to be used, and then only in its de-identified form. This distinguishes the process from "Secondary Use".

4. What type of images and information are wanted?

4.1. What type of images are wanted?

The short answer is any kind of medical image that is in digital form.

A longer answer is more complicated. The answers to the questions below describe the different types of technology, body part, disease and media that we need. The bottom line though, is that radiology and cardiology images on CD are nowadays provided as a matter of routine to doctors and patients, and any other type of image or media will be harder to obtain.

4.2. What types of technology (or "modality") are wanted?

Any type of traditional "radiology" imaging, such as X-ray (plain or angiography or fluoroscopy), CT, MRI, PET, Ultrasound, and Nuclear Medicine is useful. Nowadays, most of these types of imaging are digital, though some traditional "plain film" X-ray machines are not.

Cardiology images. Cardiac angiography in particular is now mostly performed on digital equipment, as is echocardiography and nuclear medicine exams.

Dental x-ray images.

Ophthalmology images, such as photographs of the retina, fluorescein angiography.

Dermatology images, such as photographs of the skin rashes and lesions.

Pathology images, such as micro-photographs of biopsy or surgical specimens.

4.3. What parts of the body are wanted?

Any body part is of interest, since almost any type of digital image is amenable to improvement of software and tools for viewing, analysis and interpretation.

Of particular interest though, are those parts of the body where new technology for efficiently screening healthy people for disease are being investigated. This includes images of the breast (mammography), chest (both digital x-ray and CT), heart, colon, spine and joints.

4.4. What types of disease are wanted?

Any disease is of interest. Healthy normal images are also required (see here).

Of particular interest are those diseases in which imaging plays or may play a major role in diagnosis or screening and which affect a large percentage of the population. This includes any screening or diagnostic images for cancer, for example, as well as cardiovascular disease and degenerative neurological conditions.

Also of interest are those diseases that are particularly rare yet have significant findings on imaging. Though perhaps less useful for research these may be particularly valuable for education.

4.5. Are normal images wanted?

Yes. Normal images and images of healthy people are also required.

Fortunately, a large proportion of patients who undergo imaging are normal, or have only minor findings. You should not hesitate to contribute your normal exams (or what some people refer to as "negative" exams). These are very useful for many research and testing applications, including, for example, designing tools to distinguish automatically between normal and abnormal.

4.6. What types of media can be used?

The very short answer is anything that your computer can read, since you have to get it into your computer to send to us.

The practical short answer is a CD in DICOM format, which is the standard means of distributing images throughout the radiology and cardiology world.

Normally, whenever you go to an imaging center for an x-ray or scan, they should offer you a copy of your images on CD rather than, or in addition to, printed film. If you don't get a CD as a matter of routine, see the section on how to obtain a CD for further information.

Some older equipment, though digital, cannot write CDs directly, and store images on media that by modern standards are considered weird and obsolete, not to mention expensive, such as MODs (Magneto-Optical Disks). Further, these media are often not recorded in standard DICOM format, so even if the contributor had a drive on their computer capable of reading them, they would likely not have the proprietary software. Since, in the real world, almost every imaging facility has switch over to CD, they will almost certainly have some means of converting the images to CD format.

Film cannot be accepted.

4.7. Why is film not accepted?

The short answer is because film cannot be uploaded electronically over the Internet!

When you have a sheet of x-ray or printed film, you need to hold it up to a light to see it (i.e., it is a transmissive rather than reflective medium). The range of "optical density" on the film is very broad, including very dark and very light areas, all of which are important. X-ray film is also usually physically very large (e.g., 17x14 inches). Most contributors will not have the type of scanning (digitizing) equipment necessary to scan transmissive media at all, much

less of the size and density range required. Such scanners are very expensive. Further, even if the film were to be digitized using the proper equipment, the quality is usually too low to be of use for research, which is generally focused on digital rather than digitized images. The file sizes for digitized film are very large, leading to long transfer times. Removal of the patient's identifiable information in the image makes de-identification prior to distribution considerably harder.

Fortunately, most "complex" types of imaging, such as CT or MRI, are digital in the first place. Though the images are often "printed" to film for the doctor who ordered the exam to look at or for the radiologist to interpret, by hanging the printed films on a light box, nowadays "softcopy" interpretation and distribution are used, i.e., the doctors look at the images on a computer workstation rather than printing the film. So, even if printed films are produced and handed to you, or can be, you can always request a CD, instead, or in addition; see the section on how to obtain a CD for further information.

Unfortunately, one important type of imaging is not yet universally digital. Traditional "plain film" x-ray is often still performed on older equipment that does not yet use a digital detector; rather, the x-rays are used to expose a "cassette" containing a fluorescent screen and a sheet of photographic x-ray film. This type of equipment is rapidly being replaced by so-called Digital X-ray (DX or DR) or Computed Radiography (CR) devices, which avoid these problems and produce digital images that can be recorded on CD. There is nothing "wrong" with non-digital film-screen technology from a quality or diagnostic perspective; each technology has its pros and cons. The digital technology is increasing in popularity because it allows one to save on the high cost of producing and storing and distributing (the single copy of) the films.

The bottom line for contributors of plain film x-rays is that if the facility can produce them digitally and record them on a CD for you, then great. You may need to ask them specially though, since even though their equipment is digital they may be used to routinely printing the images to film for the referring doctors.

5. What do I need to know about CDs?

5.1. How can I obtain a CD of my images?

This is discussed extensively in the introductory section for patients here.

5.2. What is a DICOM CD?

DICOM is a very common format for exchanging medical images on networks and on physical media like CDs. It defines a file format that is adopted by almost all modern medical imaging device manufacturers.

5.3. What is an IHE PDI CD?

You may also hear of the DICOM CD format being described as "IHE PDI" (Integrating the Healthcare Enterprise Portable Data for Imaging); that is essentially just another name for the same thing, and you do not need to be concerned about the difference.

5.4. Are all image CDs DICOM compliant, and how can I tell?

No. Some manufacturers refuse to provide DICOM CDs from their products. Others provide a proprietary format as an "alternative" to DICOM, may claim it is better in some way, and make it difficult for imaging centers to routinely provide DICOM CDs to patients. Sometimes sites have older equipment that has not been upgraded to generate DICOM CDs.

This problem is much less common that it used to be, since recalcitrant manufacturers have been feeling the pressure to conform. It is probably best not to worry about it, and just go ahead and upload the CD regardless, though sometimes the result will not be usable.

Though PCIR could provide software to "check" whether or not a CD is DICOM compliant, we have chosen not to, since we would rather have non-compliant or partially compliant but still usable disks uploaded, rather than have you reject them unnecessarily. It is also unlikely that even if you were to know that the CD was non-compliant, that your imaging center could (or would) do anything about it.

6. Who may contribute?

6.1. Can I contribute images on behalf of someone else?

No.

We require that the subject of the images agree to the terms and conditions during the contribution process, and currently provide no other mechanism to separate the agreement process from the contribution process.

6.2. Can I contribute my child's images?

No.

We require the contributor themselves to be able to agree to the terms and conditions and this requires them to have reached the age of maturity from a legal standpoint. The images are released to the public domain, and there is no way to revoke this. If a parent or guardian contributed on a child's behalf, the child might object on reaching maturity.

However, since it is very important for researchers to have access to images of children as well as adults, PCIR is exploring the legal and ethical considerations further.

6.3. Can my doctor or their staff contribute my images on my behalf?

No. You have to do it yourself.

Not only do we require your agreement at the time of contribution, but healthcare providers are in many countries limited by law with respect to their ability to share your information, even with your consent.

7. Access to my de-identified images.

7.1. Who will have access to my images, once they have been de-identified and released?

Anybody with access to the Internet.

The PCIR is a publicly accessible resource, and as such places no restriction on who may download the images or for what purpose. No registration is required, no fee is charged, and no "data use agreement" is required.

7.2. Can I change my mind later, and "revoke" the contribution?

No. This is stated clearly in the agreement

Once the images have been distributed publicly, there is no way to "get them back", since anyone may have downloaded them.

8. Other contribution-related questions.

8.1. Is my contribution guaranteed to be accepted and appear in the PCIR?

No. Occasionally, contributed images may not be acceptable for some reason.

Most commonly this will be caused by CD or image file format problems. For example, the CD that you receive from the radiology facility may be in a non-standard format (i.e., not DICOM) or one that the PCIR cannot recognize or cannot de-identify. For this reason, we recommend that patients explicitly request DICOM compatible CDs, but non-DICOM CDs can still be uploaded and we will do our best with them.

Sometimes the images will be damaged or be of insufficient quality. It is unlikely that these would be completely rejected, since "bad" images may have value in their own right, for example for testing the robustness and stability of software.

Images may be rejected because they are too difficult to de-identify, or an attempt has been made to de-identify them that results in excessive degradation in quality.

8.2. Will I be contacted if there is a problem with my contribution?

No.

We do not collect any of your contact information such as your email address, telephone number or mailing address. Normally, none of this information is present in the images or scanned documents either. So, PCIR is not be able to.

8.3. Will my contribution be available to the public immediately?

No.

Though much of the de-identification and indexing process is automated, there are certain manual steps that may involve a human worker, particularly those related to checking the thoroughness of de-identification and the overall image quality and completeness.

8.4. Can I make multiple contributions over time?

Yes, and we strongly encourage this.

It is very valuable to researchers to study disease progression over time, and indeed a very important area of research is the automation of change detection. For example, it one of the characteristics of a tumor is that it increases in size, and image processing to detect such size changes early is an active and promising area of development.

Accordingly, we would like you to contribute all your examinations as you have them.

In addition, it is quite likely that additional information may come to light later than at your initial contribution. For example, you may not undergo a biopsy or surgery immediately, but rather after some interval. We would like you to scan and submit such documents as pathology reports even if they are not available immediately.

If you use the same computer to perform the uploads at different times, the information stored from the previous session will help us to perform a match against the prior contribution. This is not a requirement however, and you may upload from any computer, and we will use other means to perform the match.

8.5. Can multiple people sharing the same computer contribute?

Yes.

It is preferable if they login to the computer as different users, but this is not a requirement.

9. What distinguishes the PCIR from other image repositories?

9.1. How is the PCIR different from other research image archival projects?

Some large multi-center collaborative research efforts are making the image collected during the clinical trials available in de-identified form. This is wonderful, and of tremendous benefit to the research and engineering community. Indeed, in our opinion, making one's data available publicly and openly like this should be a pre-requisite for national funding.

However, since only a small proportion of patients are actually enrolled in clinical trials (e.g., less than 5% of cancer patients is a number often quoted), and healthy normal patients are even less frequently involved, the PCIR provides an opportunity to gather information that might otherwise be "wasted".

Also, some such sites restrict access to "approved researchers", and may place non-trivial obstacles in the path of others wanting to use the data. This is often a consequence of how the images are gathered in the first place, and the limited permission obtained from the imaging subjects initially. Experience with these sites is what has led the PCIR to take the approach of seeking contributions to the "public domain", so as to be able to provide completely open access.

Some links to other image repository sites can be found here.

9.2. How is the PCIR different from teaching file distribution projects?

A common practice in medical image is to collect and distribute images for teaching purposes. These are referred to as "teaching files" and serve a vital role in training of new radiologists as well as expanding and sharing the knowledge of experienced radiologists.

With the increasing popularity of the Internet as a means of distributing information, several efforts to combine and share much larger teaching collections have been initiated. Some links to such teaching file sites can be found here.

Unfortunately, most of the teaching file cases are not useful for research of engineering purposes. The teaching file cases tend to be a very selected sub-set of images for each case, converted into a form suitable for display in a web browser (like GIF or JPEG). These images can neither be loaded into medical imaging software tools, nor analyzed by software algorithms. In short, they are neither the complete set of images nor in DICOM format.

Even though some of the teaching file repositories allow for the original DICOM files to be uploaded, it is simply not common practice for those contributing to do so.

10. Relationship of the PCIR to clinical trials?

10.1. Is contributing to the PCIR an alternative to participating in a clinical trial?

Absolutely not!

The PCIR strongly encourages you to participate in any clinical trial available to you. Clinical trials are the primary mechanism by which advances in medical science are made available.

Some links to further information about participating in clinical trials can be found here.

10.2. Images of me were acquired during a clinical trial; can I contribute them to the PCIR?

There are really two questions here; can you, and should you?

You should always be entitled to a copy of your own images. If you are physically in possession of your images, unless you are in some way contractually prohibited from doing so, there is probably no reason why you cannot contribute them to the PICR. This is especially true if the images have been acquired as part of your normal clinical care, and paid for by your normal private or national insurance or some other third party payor, rather than by those conducting the trial.

However, there are some clinical trial designs in which it may be important that the "readers" of the images in the trial are not "exposed" to the images outside the context of participating in the trial. There are also certain types of trial in which it may be desirable to sequester a "test" set of images that human readers and software algorithms have not been exposed to during "training". Whether these issues really matter in the grand scheme of things or not is contentious. Statisticians designing clinical trials have ways of dealing with such issues. The de-identified images distributed by the PCIR have no link back to you or the original clinical trial. Regardless, we suggest that you consult with the staff conducting the trial to be sure that there are no such concerns before contributing.

10.3. How can I be sure that my clinical trial images will be re-usable?

You, the patient and the subject of these clinical trials, can help directly. When asked to enroll in a clinical trial you can insist that the informed consent that you sign include a provision to make your images and other data publicly and freely and openly available for re-use immediately after the trial has finished.

Many clinical trials involve the collection of images that are never re-used beyond their original purpose. Also, sadly, many clinical trials "fail", in the sense that the outcome is not

useful to those conducting the trial (e.g., the drug does not work), and neither the results nor the data are ever published or made available.

Increasingly, national governments and others who fund clinical trials have recognized the importance of making data and results openly available, both for successful and for unsuccessful trials, and some are even predicating their funding on this. However, there are some researchers who resist this trend, for whatever reasons. Further, there are many commercial clinical trials funded and conducted by drug and device companies, who are not normally motivated to make their data openly available, although will often do so if asked.