"Storage, Backup and Archiving of Images"- E-Dermatology Task Force (IADVL Academy)

Abstract

Dermatology is a specialty that relies a lot on the visual aspect of clinical diagnosis. Taking photographs of skin lesions is a routine part of clinical practice. They are used to document skin changes during the course of treatment. It is useful for publishing articles in academic journals and textbooks, and serves as a teaching aid. Images are an important form of patient data. Like all patient data, the need for ensuring confidentiality and security is of paramount importance. Keeping all this in mind, it is important to know how to store, backup and archive your images in a safe and efficient manner.

Keywords: *Archiving, backup, clinical photography, images, storage*

Suneil Gandhi, Feroze Kaliyadan[1], Kingshuk Chatterjee[2], Aseem Sharma[3]

Consultant Dermatologist, Reviva Skin Clinic, Surat, Gujarat, [1]Department of Dermatology, Sree Narayana Institute of Medical Sciences, Ernakulam, Kerala, [2]Department of Dermatology, Calcutta School of Tropical Medicine, Kolkata, West Bengal, [3]Chief Dermatologist, Skin Saga Centre for Dermatology, Mumbai, Maharashtra, India

Address for correspondence:
Dr. Feroze Kaliyadan, Department of Dermatology, Sree Narayana Institute of Medical Sciences, Ernakulam, Kerala - 683 594, India. E-mail: ferozkal@hotmail.com

Access this article online

Website: www.idoj.in

DOI: 10.4103/idoj.idoj_642_21

Quick Response Code:

Introduction

Storage is a place where items may be accumulated or routinely used.[1] In the context of images, they are initially stored on the memory card of your camera, internal storage of your mobile phone or computer hard drive. Backup is a process of storing your items in alternate locations for the sake of safety. If you ever lost your camera or memory card, backup will ensure you have another copy of your images. Archive is a place for storing older items, maintained for occasional use/retrieval. These would be clinical photographs of patients you saw a long time ago, but have retained it for documentation or publication purposes. It can be retrieved to compare images when a patient comes to you for follow-up after a long time. These images are typically stored on hard drives and/or on cloud servers.

This article will discuss the various types of storage devices, including a few tips on how to use them effectively. The techniques for backing up and archiving your images in an efficient manner will also be discussed.

Storage

International business machines corporation (IBM) created the first commercial hard drive in 1956, which weighed over a ton and could store only 5 megabytes.[2] Apple created a hard drive for home consumers in 1981, but that could also store only 5 megabytes and was priced at a whopping 3500 US dollars.[3] We have come a long way since the days of floppy disks and compact discs (CDs). Devices are now cheaper and more data can be stored for the same cost.

There are various ways to store clinical photographs of a patient. The old-fashioned way was print them and store in photo albums. However, this is no longer in vogue, with the advent of digital images and cost-effective options for digital data storage.

Memory card

A memory card is a device that stores memory in an electronic chip, known as flash memory.[4] There are various types of memory cards, but the most commonly used ones are the SD (Secure Digital) and micro SD cards. SD cards are typically used to store images in digital cameras, whereas micro SD cards are used in mobile phones. Apart from standard capacity, newer SD and micro SD cards have the suffixes HC (high capacity) and XC (extended capacity). They basically differ in their storage capacity and speed of data transfer, SDXC/MicroSDXC having the maximum. Pen-drives and thumb-drives, called so because of their small size, also use flash memory to store data.

How to cite this article: Gandhi S, Kaliyadan F, Chatterjee K, Sharma A. "Storage, backup and archiving of images"- E-dermatology task force (IADVL academy). Indian Dermatol Online J 2022;13:321-5.

Received: 14-Oct-2021. **Revised:** 20-Nov-2021. **Accepted:** 19-Dec-2021. **Published:** 05-May-2022.

Memory cards have a write speed and a read speed. The write speed determines how fast your camera can transfer data onto the card. A higher write speed is useful if you click multiple photos in quick succession or take videos. The read speed determines how fast your card can transfer data from it to other storage devices. The higher the read speed, the better.

Always have two memory cards, one as a spare if the other fails midway during your photography session. The card should ideally be formatted in the camera you intend to use it, as every camera manufacturer have a different file storage system. Label the cards clearly and store the spares in a memory card case. Never remove a memory card from your device when it is on. Make sure that your camera battery is charged so that it does not suddenly stop in between. Stop taking further photos if an error message comes on the screen, as it can damage the memory card. Memory cards have a tiny write protect/lock switch on their left edge with the label side facing up. If turned on, this prevents accidental formatting and loss of images when the memory card is in use. This gives the name secure digital (SD) to these cards.

Hard disk drive (HDD) and solid-state drive (SSD)

A hard disk drive is an electro-mechanical storage device that has a rotating platter coated with magnetic material. A magnetic head moves above the platter to read and write data. A solid-state drive uses flash memory and integrated circuit chips to read and write data. There are no moving parts in an SSD, which reduces the risk of disk failure. Solid-state drives offer greater data transfer speeds as compared to hard disk drives. However, they are more expensive and offer lesser storage capacity for the same price as their HDD counterparts.[5]

These drives are available as internal or external storage drives. The external drives, also known as portable drives, have to be handled with care. Store them in a shock proof casing to prevent accidental bumps or scratches. Newer HDDs also come with 'active protection' by virtue of in-built accelerometers which detect abnormal vibration and acceleration. This helps unload the head, severs contact with the platter, thereby protecting the drive. Ideally, the external drive should be ejected from the computer instead of just pulling it out directly from the USB slot. Periodically check the health of your drives using computer software such as Crystaldiskinfo®, so that you may know of any potential disk failure in advance.

Network attached storage (NAS) drive

As the name suggests, a NAS drive is accessible over a network, either local or internet as opposed to 'direct attached storage '(DAS) where the computer accessing the data is directly connected to the storage device, without any network device in between. DAS tends to be faster, but for all practical purposes in the context of clinical image data, there would not be significant discernible difference between the two in terms of data transfer speed. If you are using a NAS and your mobile phone or computer is connected to the same network, you can easily transfer your images to and from that particular NAS drive. It is a device that consists of an array of two to six drive slots, integrated with its own power supply and processor. The drive slots can be filled by HDDs or SSDs, depending on the model. The multiple drive slots can be used to upgrade your data storage capacity. You can also create copies of your files in multiple drives as a backup in case one of the drives ever fails. It can also be programmed to create backup copies at specified intervals. For all these reasons, NAS drives are considered to be an elegant solution for private cloud storage.[6]

Cloud storage

This is an off-site form of data storage that is maintained by a third party. It safely and securely stores your data in remote data servers. This can be easily accessed via a website or app over the internet, otherwise known as the 'cloud'. You can streamline your storage by uploading files and accessing it from multiple devices. It is easier to share files. You can feel secure, knowing that all your important data is backed up in a safe, offsite location. This is especially useful in times of disaster recovery if your hard drive fails for any reason.[7]

The minimum bandwidth recommended for cloud computing is 80 kbps. With the widespread penetration of high speed internet, bandwidth should not be a major limitation for switching to cloud storage. However, depending on the network and location, accessing cloud data might be difficult. This needs to be factored in while deciding on storage. If the data you are storing needs to be retrieved frequently, it may not be a good idea to rely on the cloud as the primary data storage point. On the other hand, if you need to access data from different locations (for example you work in multiple centers), then cloud storage is indeed the best option. There are many cloud storage providers available in the market today. We are most familiar with the free storage that comes inbuilt with Google, Microsoft and Apple services. However, there are a few drawbacks to free cloud storage, listed in Table 1.[8] You will eventually have to move to paid cloud storage options if you require to store more data, with scaling up of clinical practice. Electronic medical record (EMR) software service providers offer the option of cloud storage included in their plans at affordable rates. They usually use Amazon Web Services (AWS)®, Google Cloud®, Backblaze® or Microsoft Azure® as their cloud service. One disadvantage of using cloud storage is the need to depend on a third party to host your data. Therefore, while using a cloud service, either for images alone or as part of EMR, it is important that you read the fine print on data confidentiality and security. Most of the major services

Table 1: Differences between free and paid cloud storage

Free cloud storage	Paid cloud storage
Limited capacity (usually <15 GB)	Higher capacity (100 GB - 2 TB+) Some plans are unlimited or pay as you use.
Less privacy (companies may see your data, may use them to target ads)	More privacy (zero logging - company cannot see your data)
Less secure	More secure (file encryption, 2-factor authentication, etc.)
Lesser features	More features (added security, file versioning, etc.)
Ads may be displayed	Ad free
Higher chance of downgrading or limiting services in the future	Less likely to be downgraded

offer good security and encryption (as in the case of other cloud services like emails, for example). In the context of teledermatology consults, remember to keep your image files backed up on the cloud. For messaging services like WhatsApp, there is an automatic back-up option where you can set the frequency of back-up. It is important to save a copy of the back-up files to your PC or hard-disk to ensure that you will not lose files inadvertently while deleting back-up to free space.

Backup

Most of you would have faced the frustration of losing important files at some point of time in life. The 3-2-1 rule is a best practice method for backup of files.[9]

It means that you should:

Store at least three copies of your data. This includes the original copy and at least two backups.

Use two different types of storage formats. The chances of two entirely different storage formats failing, or becoming incompatible in the future are less. Therefore, if you have data stored on an internal hard drive, make sure you have a secondary storage type, such as external hard drive or cloud storage.

Keep at least one copy of the data offsite. Even if you have two copies on two separate storage types, but both are stored in the same location, you could lose both of them to a natural disaster or theft. Keep a third copy in an offsite location like the cloud or at other premises.

The 3-2-1 rule guarantees that you will always have a copy of your data. Multiple locations ensure that there is no single point of failure and that your data is protected from natural disasters like fire and floods.

You must periodically back up your images. This can be done either as a full, incremental or differential backup.[10] A full backup, as the name implies, makes a copy of your entire data onto a selected storage. If the size of the backup is large, it can take a long time to copy everything. This makes it impractical to do so on a regular basis.

Incremental and differential backups are a quicker way to backup data after one full backup has been created initially. Incremental backup only backs up data that is different from the previous backup. Every incremental backup is stored as a separate file set. The sets will be smaller in size and therefore faster to backup. The drawback to this method is that each incremental backup depends on all the file sets before it. Even if one set is corrupted, the recovery process will be affected. Recovery of the entire data file set is slow, as multiple incremental sets have to be combined.

A differential backup is slightly different from an incremental backup. It backs up data that has changed since the last full backup in a single file set. Recovery time of the entire data file set is comparatively faster than incremental backup, as only two file sets have to be combined. The various types of backups are illustrated in Figure 1.

Full backups can be done once a month, while incremental/differential backups can be done on a daily or weekly basis. Multiple copies of your data must be stored as discussed in the 3-2-1 rule of backup. This will ensure that even if you have accidentally lost or deleted files, you can easily retrieve them from your other copies. Recuva® is a useful computer software that helps to retrieve accidentally deleted files. However, this works only if the storage has not been overwritten with other files.

Archiving

A few years into your career, you will have a sufficiently large number of images. This will not be accessed regularly. Archiving these images in a systematic manner will make it easier to retrieve them later.

The images should not be stored on the same drive on which the computer operating system is installed. In case of a software problem or malware attack, the operating system may need to be reinstalled. The entire drive may need to be formatted, which results in erasure of all the data on it.

The images should be placed in a folder that is clearly labelled. Separate subfolders can be created to organize the images by year, month, disease name and patient ID. The labelling structure should be consistent. One example of a drive path with a subfolder structure would be D:/Clinical Photographs/2021/September/Psoriasis/ABC-123. Doing so, one can easily search for specific image folders belonging to the same patient. One can also search for all image folders related to a particular disease.

Keywords, also known as tags, are an excellent way to have more granular control on accessing clinical

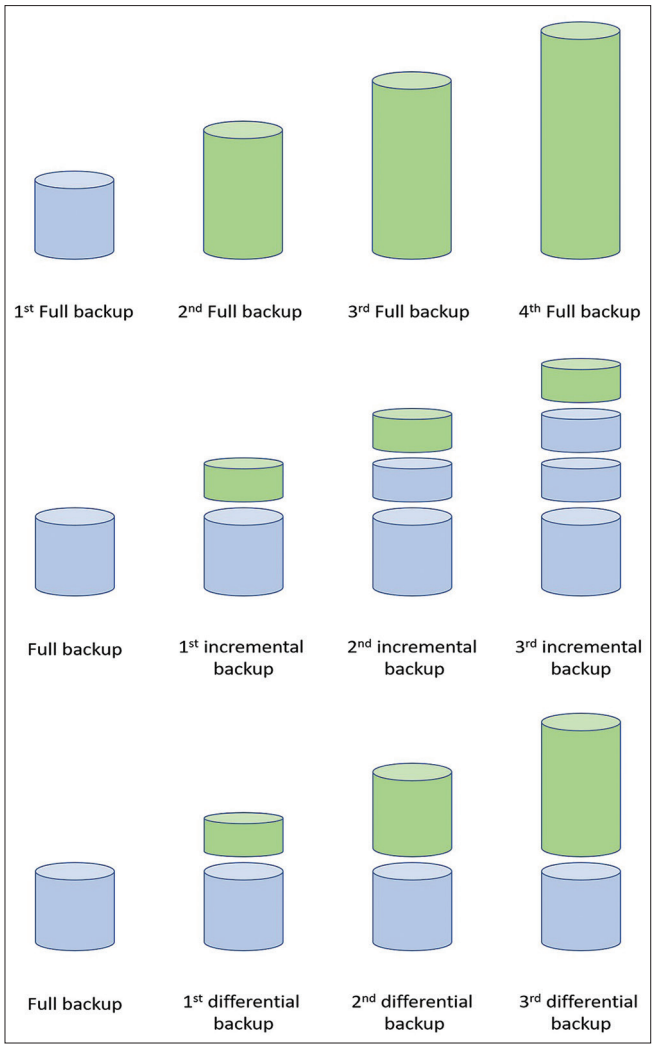

Figure 1: Full, incremental and differential backups. The green cells represent the amount of data copied in each backup

photographs. Image tags are especially helpful in retrieving series or groups of related images. This could be for pre-post comparisons, to assess therapeutic response to medical or surgical treatment, or for research studies. In addition to the folder structure organization, keywords can be added to each image file separately. The disease name, regional location, grade/stage/classification, or type of procedure/treatment modality are examples of keywords. Pre-treatment and post-treatment are useful keywords that help to serially track response to a particular procedure/treatment modality. There is no limit to the number of keywords that can be used. For example, the keywords 'psoriasis, nail, intralesional, methotrexate' can be added to the image of a nail psoriasis patient that was treated with intralesional methotrexate. Once keywords are added to all your images, you can quickly retrieve all images of patients suffering from psoriasis by typing the keyword 'psoriasis'. The search can be narrowed down to all patients of nail psoriasis by typing the keywords 'nail, psoriasis' in the search bar. You can quickly locate images of all patients you treated with intralesional injections, regardless of what disease they have. With the smart use of keywords, the possibilities are endless.

Archiving is essential from multiple points of view. Legally, all data including images are owned by the patient. They have the right to request for the same in the 'as-is' original format. Hence it is important to not only archive properly, but also in a manner that is easy to retrieve as and when needed.

The other important facet of archiving is in research. Research studies extending over long periods can generate a large amount of image data, especially in the field of dermatology. Proper archiving is essential to ensure that the correct images are available and easily retrievable as and when needed. For research related images, we recommend to use copies of the original images in a separate folder specifically for the research files (without altering the labeling/filing system for the original images). This ensures that the original files are available in case the research files are lost for any reason. For studies needing serial images, it is important to follow the labeling for the baseline image and add a sequence number and date for each image (as long as the original image is archived, the date can be retrieved from the meta-data, but adding the date to the label itself makes retrieval and arrangement much more easier. While adding the date to labels, it is important to follow a standard format throughout (YYYYMMDD is best for retrieval purposes).

In the context of larger clinics, the most important aspect is to ensure that all members of the team are aware of and follow a standardized protocol for storage and archiving. In clinics with electronic health records, adding the medical records number to the labeling ensures that in case of any confusion related to patient details, the same can be cross-checked easily, irrespective of who is seeing the patient at a particular time. In clinics with multiple team members, it is important be clear on access and editing privileges available to each member of the team. Most Electronic Health Record/Electronic Medical Records (EHR/EMR) systems now have options to add images to the patient records. While it may be difficult and tedious to add a large number of images for each visit, it would be good to add at least the most important images. The images linked to the patient record should be unedited images in the original form.

Privacy and security of images

It is important to note that images of patients are sensitive information. The goal is to protect patient privacy, keeping identifying information private and secure. Unfortunately, data breaches are becoming more common. Medical images are most vulnerable. A global medical data breach in 2019 exposed records of more than 120 million Indian

patients.[11] This contained images of X-rays, MRIs, CT scans and patient photographs. Safe practices to protect your images include file encryption, strong passwords and two-factor authentication. Only authorized people must be allowed to access your files and devices. A periodic storage and network audit must be performed to verify their integrity.

The laws around data privacy and security are very strict in developed countries. The European Union has GDPR (General Data Protection Regulation), while USA has HIPAA (Health Insurance Portability and Accountability Act). Hefty fines are imposed for non-compliance. The Indian IT act is outdated, but bills such as the PDP (Personal Data Protection) bill and DISHA (Digital Information in Security Healthcare Act) are being discussed in the parliament and will be implemented in the near future.[12]

Teledermatology has received a renewed impetus due to the COVID-19 pandemic. It is mandatory to store all electronic health records, including relevant images during the lifetime of the patient for medicolegal purposes (and preferable to preserve indefinitely). The patient data is basically 'owned' by the patient. He/she has the right to request for it any stage, and this includes images. Moreover, the 'as-is' principle applies - data captured, should be available in the original format. Any form of electronic data, including images, should be subject to stringent security to ensure that patient confidentiality is not compromised at any stage.[13] Therefore, it is important to know how to store them in a safe and efficient way. A suggested workflow is depicted in Figure 2.

Conclusion

Image storage, backup and archiving has come a long way from the days of film. File formats have changed, storage media have upgraded and cloud storage has come into the fold. Clinical images are an inevitable aspect in the daily life of a dermatologist. Like all kinds of patient data, they need to be stored in an effective and secure manner. A combination of hard disk storage and cloud storage is the best way to store them. Simple techniques like proper keywording/tagging of images can help in easier retrieval. Moreover, device backups must be secured, and drive backups must be reinforced. Even cloud storage must have a surrogate. Nothing lasts forever, but data surely can.

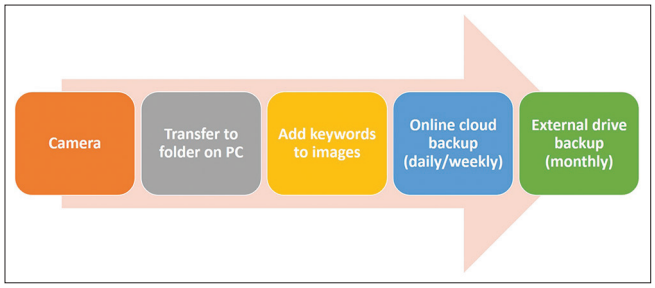

Figure 2: Suggested workflow of storage, backup and archiving images

Acknowledgements

We would like to acknowledge all the e-dermatology task force members and the IADVL academy, for the valuable inputs.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

References

1. Store vs Archive - What's the difference?. Available from: https://wikidiff.com/archive/store. [Last accessed on 2021 Sep 30].
2. Computer history museum. 1956: First commercial hard disk drive shipped. Available from: https://www.computerhistory.org/storageengine/first-commercial-hard-disk-drive-shipped/. [Last accessed on 2021 Sep 30].
3. Apple ProFile. Available from: https://en.wikipedia.org/wiki/Apple_ProFile. [Last accessed on 2021 Sep 30].
4. Memory card. Available from: https://en.wikipedia.org/wiki/Memory_card. [Last accessed on 2021 Sep 30].
5. SSD vs. HDD: Which is better for you?. Available from: https://www.crucial.com/articles/about-ssd/ssd-vs-hdd. [Last accessed on 2021 Sep 30].
6. Brooks T. What Is a NAS (Network Attached Storage)? [updated 2021 Aug 13]. Available from: https://www.howtogeek.com/742893/what-is-a-nas/. [Last accessed on 2021 Sep 30].
7. Powerful cloud storage for all your files. Available from: https://www.dropbox.com/features/cloud-storage. [Last accessed on 2021 Sep 30].
8. Is Free Cloud Storage Any Good?. Available from: https://www.cloudstoragereviews.co/cloud-storage-basics/is-free-cloud-storage-any-good.html. [Last accessed on 2021 Sep 30].
9. Vanover R. What is the 3-2-1 backup rule? [updated 2021 Jul 20]. Available from: https://www.veeam.com/blog/321-backup-rule.html. [Last accessed on 2021 Sep 30].
10. Full vs Incremental vs Differential Backup: A Detailed Comparison [updated 2020 Mar 30]. Available from: https://phoenixnap.com/kb/full-vs-incremental-vs-differential-backup. [Last accessed on 2021 Sep 30].
11. 120 Million Medical Records Leaked! Global Medical Report Sheds More Light. [updated 2020 Feb 06]. Available from: https://www.ehackingnews.com/2020/02/120-million-medical-records-leaked.html. [Last accessed on 2021 Sep 30].
12. Luniya V. India: DISHA – India's Probable Response To The Law On Protection Of Digital Health Data [updated 2021 Apr 20]. Available from: https://www.mondaq.com/india/healthcare/1059266/disha-india39s-probable-response-to-the-law-on-protection-of-digital-health-data. [Last accessed on 2021 Sep 30].
13. e-Health division, ministry of health and family welfare, government of India. Electronic Health Record (EHR) Standards for India 2016. Available from: https://main.mohfw.gov.in/sites/default/files/EMR-EHR_Standards_for_India_as_notified_by_MOHFW_2016_0.pdf. [Last accessed on 2021 Sep 30].