Current and future accelerator technologies for charged particle therapy

DOI:
10.1016/j.nima.2015.08.038

Document Version
Accepted author manuscript

Link to publication record in Manchester Research Explorer

Citation for published version (APA):

Published in:
Nuclear Instruments & Methods in Physics Research. Section A: Accelerators, Spectrometers, Detectors, and Associated Equipment

Citing this paper
Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights
Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy
If you believe that this document breaches copyright please refer to the University of Manchester’s Takedown Procedures [http://man.ac.uk/04Y6Bo] or contact uml.scholarlycommunications@manchester.ac.uk providing relevant details, so we can investigate your claim.
Current and Future Accelerator Technologies for Charged Particle Therapy

Hywel Owen\textsuperscript{a,b,*}, Antony Lomax\textsuperscript{c,e}, Simon Jolly\textsuperscript{d}

\textsuperscript{a}School of Physics and Astronomy, University of Manchester, UK
\textsuperscript{b}Cockcroft Institute for Accelerator Science and Technology, Daresbury Science and Innovation Campus, Warrington WA4 4AD, UK
\textsuperscript{c}Paul Scherrer Institute, Villigen, Switzerland
\textsuperscript{d}Dept. of Physics and Astronomy, University College London, UK
\textsuperscript{e}Dept. of Physics, ETH Zurich, Switzerland

Abstract

The past few years have seen significant developments both of the technologies available for proton and other charged particle therapies, and of the number and spread of therapy centres. In this review we give an overview of these technology developments, and outline the principal challenges and opportunities we see as important in the next decade. Notable amongst these is the ever-increasing use of superconductivity both in particle sources and for treatment delivery, which is likely to greatly increase the accessibility of charged particle therapy treatments to hospital centres worldwide.

Keywords: Keywords

1. Hadron therapy

The development of modern physics and modern medicine have gone hand in hand, and the end of the 19th century heralded not only the birth of the atomic age but also of the field of radiotherapy. Only a few months after the discovery of radioactivity and x-rays \cite{1}, high-energy photons were already being used to treat diseases of the skin \cite{2,3,4,5} and shortly thereafter to treat cancer \cite{6}. Today, treatment with x-ray photons remains the mainstay of radiotherapy and...
has become — with surgery and chemotherapy — one of the three pillars of cancer treatment. Modern radiotherapy devices have achieved a remarkable degree of sophistication, and complex techniques such as intensity-modulated radiotherapy (IMRT) are now commonplace in the hospitals of the developed world [7, 8, 9, 10, 11, 12, 13].

Whilst treatment with photons is still predominant, already in 1947 Robert Wilson pointed out that heavy charged particles such as protons could have substantial advantages when delivering a therapy dose to a patient [14]. After an initial build-up distance the dose from photons decreases more-or-less exponentially with depth in the patient; a treated region deep in the body thereby incurs an unwanted larger dose proximal (upstream) to that region, and techniques such as IMRT serve to alleviate that problem. Protons in contrast mostly deposit their energy gradually; moreover the rate of energy deposition is roughly inversely proportional to that energy, which gives rise to the characteristic Bragg peak such that most of the initial proton kinetic energy is deposited close to the end of the proton’s range. A concentration of the dose deposition gives the potential to irradiate desired regions (such as tumour volumes) whilst sparing as much as possible the surrounding tissue and organs at risk (OARs).

The Bethe-Bloch equation [15, 16, 17] describes the behaviour of the energy loss as

\[
\frac{dE}{dx} = k z^2 Z A \frac{1}{\beta^2} \left[ \frac{1}{2} \ln \frac{2m_ec^2 \beta^2 \gamma^2 T_{max}}{I^2} - \beta^2 - \frac{\delta(\beta \gamma)}{2} \right],
\]

where \(Z\) and \(A\) are the atomic number and atomic mass of the absorber, \(z\) is the charge of the incident particle, \(I \approx 11.5Z\) eV is the mean ionisation potential, \(T_{max} = 2m_ec^2 \beta^2 \gamma^2 / [1 + 2\gamma m_e/M + (m_e/M)^2]\) is the maximum kinetic energy that may be imparted to an electron in a single collision, \(k = 4\pi N_A r_e^2 m_e c^2\), and \(\beta\) and \(\gamma\) are the conventional relativistic factors. \(\delta(\beta \gamma)\) is a density correction term [18]. This equation highlights many of the factors influencing the use of heavy charged particles for therapy. Firstly, the depth at which the Bragg peak occurs is set by choosing an appropriate initial energy; as an example, ICRU49 defines that a 230 MeV initial proton energy has a mean range of about 33 cm in
water, a reference material quite close to densities encountered in tissue. Higher-density materials have a greater stopping power, and heavier ions such as $^{12}$C$^{6+}$ require much greater energies up to 450 MeV/u to obtain the same 33 cm range. Uncertainties in each of these parameters — the value and spread of the initial energies, the composition and density of the material passed through, and the validity of the model used to estimate the slowing rate — are all important factors in predicting the range of dose deposition mechanisms of the incident particles. Both accurate imaging and accurate estimation of dose deposition are required to obtain the dose advantages that the Bragg peak offers.

1.1. Clinical treatment requirements

Modern radiotherapy has developed a typical regimen wherein a curative treatment divides application of irradiation to a site into several fractions spaced around a day apart. For example, a prescribed prostate treatment might demand a total treatment dose of 70 Gy to the treatment site (the planning treatment volume, or PTV) administered in 35 fractions of 2 Gy each over 7 weeks. Fractionation benefits from the nonlinear and differing response of diseased and healthy tissues to dose, whilst also spreading the dose over time to treat different phases in each cell’s cycle. Whilst there is uncertainty as to whether dose administration in very short times ($\ll 1$ s) might have a differing biological effect than if spread out over a longer time, such differences are likely to be smaller than other uncertainties such as the conversion from dose to response for a particular particle species, energy or tissue type. Indeed, the same absorbed dose in Gy gives rise to differing radiobiological effect (RBE) accounted for in planning through a factor. For protons the RBE is presently taken to be 1.1 until there are better systematic measurements; the reference RBE is taken to be 1 for the c.1.3 MeV gamma rays from $^{60}$Co decay. Heavier ions and other species such as pions can give rise to a greater RBE due in part to their higher linear energy transfer (LET). However, their higher mass and charge means that substantially higher incident energies are required for the same treatment depths in patients, necessitating larger accelerator equipment;
this is discussed later. Another issue with certain species, such as pions, antiprotons and $^{12}\text{C}$ ions, is that the sharp distal fall of dose beyond the Bragg peak may be somewhat spoilt by other processes such as annihilation or fragmentation; this effect must be balanced against the RBE advantages for some tumour sites. For example $^{12}\text{C}$ might have an RBE as much as 3.5 depending upon the cell type and end point studied, and carbon-ion therapy is indicated in particular for some radioresistant tumours [28, 29, 30, 31]. The use of heavy ions such as carbon may allow fewer fractions to be utilised; consideration of the differing biological effects is described for example in Tsujii et al. [32]. The relative merits of using protons or carbon ions for radiotherapy is the subject of some debate at present [33, 34, 35].

To determine the requirements of the accelerator one must also consider the total volume of the PTV into which the dose must be administered; the larger the volume, the larger the number of particles that must be delivered. A typical volume useful for specifying accelerators is to assume a litre-sized PTV some depth within the patient. If the depth range of the PTV is, say, 10 to 20 cm then it may be covered by varying the incident proton energy such that the Bragg peaks for each energy occur at different depths. This is the concept of the spread-out Bragg peak (SOBP), and may be achieved either by interposing a suitable varying-thickness modulator wheel [36], or by adjusting the proton energy in discrete steps by some other means; the latter method forms part of the burgeoning method of spot scanning which is replacing the previous passive spreading methods as a means of efficiently covering the PTV.

To cover this canonical 10 to 20 cm depth in water requires proton energies from about 112 to 166 MeV, and 1 Gy of dose in this litre volume requires around 90 billion protons (90 Gp) of 22 pJ average energy, or about 16 nC of charge [37]. Appropriate intensity weighting of the different-energy spots allows a homogenous dose to be obtained within the PTV; this technique is known as Intensity-Modulated Proton Therapy (IMPT). It should be noted that it typically takes some minutes to adequately settle and position a patient so that an accurate treatment may be administered, and so from the point of view of
patient throughput it is adequate to deliver that 1 Gy in about a minute. A very high proton fluence might allow tumours to be treated without correcting for patient motion, but this would be quite difficult owing to the requirement to monitor fluctuations in the delivery of those protons. Delivering 1 Gy in a minute translates to a required average proton current of less than 1 nA, a rather modest value compared to other applications of such particle accelerators; somewhat lower currents are required e.g. for therapy with carbon ions. Proton and carbon-ion treatments already benefit from motion compensation methods using techniques such as beam gating.\textsuperscript{[38, 39]}

As well as the current and energy requirements one must also consider the requirements on the transverse beam size and energy spread. As any heavy charged species such as a proton traverses a solid material, it scatters and loses energy with some statistical fluctuations. The primary ionisation energy loss gives rise to a straggling of the range; the Bragg peak has a finite width which to a good approximation is about 1.1% of the range, in other words about 1 to 2 mm. There is little point therefore in having a very small initial energy spread as it will be smeared out by the straggling, and so 0.4% energy spread is generally adequate. Similarly, multiple coulomb scattering (MCS) will spread the protons out transversely, and again an initial incident spot size of a few millimetres is adequate; this translates to an emittance requirement of around 10 mm-mrad. These beam specifications are readily achieved by conventional accelerator technology as will be seen below.\textsuperscript{[40, 37, 41]}

The final requirement is to be able to vary the direction from which the administered dose is delivered to the patient. In a very limited set of cases the patients themselves may be moved somewhat, but in most treatments the patient is supine and the incident particles must be directed from a variety of angles; this is done primarily to avoid dose to critical structures but also to spread out the proximal (upstream) dose. In this regard the gantries used in charged particle therapy serve a similar purpose to the smaller rotating electron linacs that deliver photons for intensity-modulated radiotherapy (IMRT). The most-used method of achieving a varying field angle is to utilise a so-called
isocentric gantry, where the beam is rotated around a stationary patient; this is shown schematically in Figure 1. Other geometric arrangements are possible, notably the alternative of mounting the accelerator itself upon the gantry, but the isocentric gantry is the method used in most proton treatment rooms today. Whilst a full 360 degrees of rotation means the patient would not have to rotate, recently it has also been proposed to combine a narrower range of gantry rotation angles (say, around 180 degrees) with patient table rotation; the rationale for this is to restrict the floor area required for each treatment room. Gantries may also be used for other ions, but as those species become heavier the magnetic fields required to deflect them increase accordingly. The stiffness of the ion beam is described by its magnetic rigidity $B\rho = p/q$, where the particle momentum $p$ and charge $q$ determine the trade-off between magnetic field $B$ and resultant bending radius $\rho$. For example, protons of kinetic energy 250 MeV are sufficient to treat adult patients and have a magnetic rigidity of 2.43 Tm; 1.8 T is about the largest field available from a resistive (normal-conducting) magnet, leading to a bending radius of 1.35 m. The only carbon-ion gantry currently in operation is at the Heidelberg Ion Beam Therapy Centre (HIT), where the maximum carbon-ion energy of 425 MeV/u corresponds to a magnetic rigidity of 6.57 Tm; the bend radius in their normal-conducting magnets is thus nearly three times as large (3.65 m), and to make it smaller requires the use of superconducting magnets that may deliver a larger magnetic field.

Regardless of the type of gantry, in IMPT the incident spots must be scanned transversely across the patient with good positional reproducibility. A typical transverse field size will be greater than 20x20 cm, and each spot must be delivered to the isocentre with an accuracy of 1 mm or better; to this must be added the errors from patient positioning, which are typically a few millimetres depending on the treatment site.
Figure 1: Schematic illustration of a charged particle therapy facility with isocentric gantry. The gantry consists of a number of dipole and quadrupole magnets that all rotate around the patient; in this example the magnets that scan the spots across the treatment field are located downstream of the final dipole, which is often but not always what is implemented.

2. Requirements of treatment facilities and accelerator technologies

A treatment facility consists of an accelerator source and at least one treatment room, and in most implementations today several treatment rooms are fed from a single accelerator source, physically connected by a beam transport system (BTS) that can direct the protons or ions into each treatment room. The historical reason for this was the size and mass of the accelerator sources, which originally were either normal-conducting cyclotrons or synchrotrons neither of which could be mounted on a gantry. Also, since a treatment room only need take beam for a small fraction of the time the patient is there — most of the time is taken up with patient positioning and verification — it is suitably efficient to switch a single accelerator into multiple treatment rooms. It has been shown in several modelling studies that this switching remains efficient with up to three or four treatment rooms, and the capacity of a three-room proton treatment centre can be as much as 750 patients per year. Single-room systems
may achieve about 200-250 patients per year, although this number is quite dependent on the complexity of the treatments and on whether hypofractionation (delivery of the treatment in relatively fewer fractions) is undertaken. Beam switching from room to room may take as little as a few seconds (depending on the system used), and reducing this time is seen as an avenue for improving overall patient throughput in a multiple-room facility [42, 43, 44].

2.1. Cyclotrons

Despite being the first cyclic accelerator (invented in the early 1930s), the cyclotron remains the efficient workhorse of many accelerator applications including for proton therapy [45, 37, 41]. At lower kinetic energies (less than 20 MeV) the classical Lawrence cyclotron principle may be used in which an alternating Dee voltage progressively accelerates the protons to a desired extraction energy whilst they circulate in a simple, single static dipolar magnetic field (see Figure 2). The revolution frequency is \( f = \frac{qB}{2\pi\gamma m_p} \), so at the larger kinetic energies needed for proton therapy (\( \gamma = 1.267 \) at 250 MeV) the frequency change is sufficiently large that it must be compensated either by increasing the field with radius (isochronous cyclotron) or by changing the Dee frequency (synchrocyclotron); the synchrocyclotron method is possible because of the low requirement on average current. Cyclotrons become progressively more difficult as their extraction energy is increased, and are generally designed to deliver a single extracted energy; synchrocyclotrons have a simpler magnetic design and deliver higher energies with greater ease, but have a more limited duty factor around 1000 cycles per second similar to FFAGs (see below). The fixed extraction energy must be degraded before impinging on the patient, and this is performed usually with a pair of absorbing wedges (typically graphite) that reduce the energy without undue scattering and straggling [46]; a portion of this degraded beam is then selected with collimators, and can be as little as 0.1% of the initial beam intensity at the lowest energies of about 70 MeV. To make up for this the extracted current from the cyclotron must be correspondingly higher, made up of a combination of high circulating current (up to
1000 nA) and a good extraction efficiency (usually in the range from 30% to 80%). Most existing treatment facilities utilise a single energy selection system (ESS) and then adjust the downstream BTS magnets, but more recently it has been proposed (e.g. by ProNova) to utilise a separate ESS for each treatment room; this would allow less tuning of the BTS and the utilisation of low-power permanent magnets.

Normal-conducting cyclotrons such as the IBA C235 (which has a central field of 1.7 T) and its derived Sumitomo version are the mainstay of current proton therapy accounting for around half of all treatments; these have masses around 220 tons and diameters of about 4 m [41]. However, a key trend in cyclotrons is the progressive adoption of larger fields using superconducting technology [47, 48, 49, 50, 51, 52, 53]. The first widespread example of this is the Varian ProBeam in which the central cyclotron field is 2.4 T, allowing the size of the cyclotron to reduce substantially (total mass around 80 tons) [54]: note that cyclotron mass scales roughly as $\sim 1/B^3$, so even modest increases in field can yield large savings in mechanical engineering. More recently Mevion have implemented a 9 T, 250 MeV synchrocyclotron (Figure 3) that achieves a mass of only 20 tons, and this allows it to be mounted directly upon the treatment gantry. As such the Mevion was the first true single-room solution, and in concept borrowed significantly from pioneering work on such cyclotrons carried out at MIT [55, 56]. Cyclotron companies such as IBA and Varian are putting forward single-room solutions in which the accelerator is floor-mounted but where there is only a single treatment room. The rationale for single-room systems is that the initial capital cost of a single-room system (figures of c.$40M are often publicised) is lower than for a multi-room system, but of course the capacity of one room will be proportionately lower. Having dedicated accelerators for each treatment room is also argued to give greater reliability, but this has not yet been clearly demonstrated.

Whilst 250 MeV proton energy is accessible from cyclotrons at ever-decreasing cost, they struggle to achieve energies more than this. Whilst experimental (nuclear physics-led) cyclotrons like those at PSI and Gatchina [57] have achieved
higher energies than 250 MeV, these accelerators are by no means compact and are unlikely to be suitable for widespread hospital use. These higher energies can find use particularly for diagnostic applications such as proton computed tomography. Similarly, whilst some significant design work has been performed to construct a superconducting cyclotron suitable for carbon therapy (for example the IBA C400 [51]), there is as yet no commercial product available.

![Illustration of the operation of a cyclotron](image)

**Figure 2:** Illustration of the operation of a cyclotron, here showing a classical cyclotron in which the field does not vary with radius. The generated dipole field $B$ is static, and within it the protons are gradually accelerated (after injection from an ion source) within the vacuum vessel due to the alternating voltage applied across the gap between the Dees, which may be as much as 100 kV per turn or more. FFAGs are similar except that there are multiple, strong-focusing dipoles and that the orbit radius in FFAGs typically varies much less than it does in a cyclotron.

### 2.2. Synchrotrons

The synchrotron is the other widely-used technology for charged particle therapy [58], and indeed the first hospital-based centre at Loma Linda (LLUMC) utilised a Fermilab-designed synchrotron [59, 60, 61, 62] now commercialised by Optivus. Synchrotrons differ from cyclotrons in that the magnetic guide field varies as the particles are accelerated such that the bend radius (and hence overall path) of the particles remains constant (see Figure 4): the revolution frequency and hence accelerating cavity frequency also increase through the
acceleration cycle. This results in the ability to have much smaller magnets than in a cyclotron, albeit now more are needed (in particular, quadrupoles) to provide the necessary periodic strong focusing [37]. Since the magnetic field varies there can only be one pulse of - say - protons circulating in each acceleration cycle, in contrast to the isochronous cyclotron where many different bunches are simultaneously present with different energies. To overcome this limit many more protons are placed in each accelerated bunch, as much as 10 nC or more. The acceleration cycle time is limited by inductive effects in the synchrotron magnets, and is typically chosen to be several seconds long. After acceleration the dipole field (and hence energy) is held constant — the so-called flat top — so that the protons may be slowly extracted and directed to the treatment room. Protons may be progressively extracted over several seconds such that they may also be scanned at the patient in a similar manner to the near-continuous beams from cyclotrons. The principle advantage of the synchrotron is that because the energy may be varied with each acceleration cycle there is no need for an energy degrader, which greatly reduces
the average current requirement. This, coupled with the higher proton bunch charge results in a treatment time at the patient which is much the same as from a cyclotron, that is about a minute. The other advantage is that synchrotrons may readily accelerate particles with larger magnetic rigidities than is readily achieved from a cyclotron, so at present synchrotrons are the only accelerators used for carbon therapy; present examples include HIT, CNAO, MedAustron and NIRS [37, 41, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72]. More recently, a magnetic cycle has been demonstrated (by NIRS and Hitachi) in which the dipole field and extracted particle energies reduce during the flat top citeiwata10,iwata2013,noda2014; this allows a finer granularity of each treatment layer in the patient.

Synchrotrons are a well-developed technology, and in contrast to cyclotrons it is rather difficult to adopt superconducting technology since it is difficult to rapidly vary the field of a superconducting magnet. Hence it is difficult to significantly reduce the footprint of a synchrotron other than to reduce the number of magnetic cells as much as possible, as has been done by ProTom and Hitachi. Another route for improvement is to achieve higher cycling rates, and the most notable project is the BEST Medical/Brookhaven Laboratory collaboration to develop the ion Rapid Cycling Medical Synchrotron (iRCMS) [73, 74, 75, 76]. Here the cycling frequency will be between 15 and 30 Hz, improving the ability to perform spot scanning and offering the ability in principle to treat with both protons and carbon ions in a single treatment fraction.

### 2.3. FFAGs

A compromise between cyclotrons and synchrotrons is the fixed-field alternating-gradient accelerator (FFAG) [77, 78, 79]. The FFAG is a strong-focusing variant of the cyclotron first proposed in the 1950s, and retains the advantages of having a fixed dipole field (so that superconducting magnets are an option) whilst giving a greater energy reach than is possible with a cyclotron. Initially neglected because they are dynamically more complex than either cyclotrons and synchrotrons — and so are difficult to design — they have received increased
Figure 4: Illustration of the design and operation of a synchrotron. Synchrotrons consist of a number of identical magnetic cells (shown top left) each comprised of at least one dipole and sufficient quadrupoles to obtain stable transverse focusing (shown top right). Phase stability ensures that the accelerating protons keep step with the rising dipole field such that their average orbit is fixed over the acceleration cycle (shown below). Extraction is obtained during the flat-top by one of several resonant extraction methods.

attention in the last few years with the demonstration of the first non-scaling FFAG, EMMA [80]. Since then several design studies have presented conceptual solutions that address needs unmet by cyclotrons and synchrotrons. As well as the greater energy reach (the PAMELA outer FFAG may accelerate carbon ions to 430 MeV/u [81] and the NORMA FFAG may accelerate protons to 350 MeV [82]) it is proposed to conduct fast injection and extraction from such machines [37]. With sufficient accelerating voltage the fixed field in FFAGs allows acceleration to be undertaken in a 1 ms cycle time but with a larger bunch charge than in a cyclotron. Pulse-by-pulse energy variation at 1 kHz is proposed to be achieved with appropriate fast extraction kickers, and similar concepts are now being explored to enable fast energy variation from a cyclotron. However, as yet only a partial demonstration of this sort of extraction has been conducted at KURRI and at EMMA, and as yet no clinical facility has been constructed.
2.4. Linacs

Linear accelerators (linacs) historically pre-date all the accelerator types mentioned so far [58], Wideroe having first proposed the concept in the 1920s to overcome the limitations of static voltages [83]. Until very recently linacs have not been favoured because (for protons at least) the accelerating gradient possible was limited to only a few MeV/m such that a significant energy reach demanded a long accelerator. Indeed, it was this very limitation that inspired the invention of the cyclotron and synchrotron, and linacs until recently were only used at lower energies (as injectors into synchrotrons for instance) or in situations where very high currents are required (which is not the case in charged particle therapy of course). These limits have long been recognised, but in recent years there have been significant developments toward achieving the gradients needed to make linac-based therapy systems size- and cost-competitive with other technologies. For example, both the TERA and TOP-IMPLART projects have demonstrated 3 GHz accelerating structures capable of greater than 20 MV/m [84, 85, 86, 87, 88], and the TERA LIBO side-coupled structure has been adopted as the technology for the Advanced Oncotherapy (AVO) linac-based solution presently under construction at CERN. Similar to other applications of linacs one issue is about beam injection into such high-frequency structures. The conventional approach is to utilise an ion source, radio-frequency quadrupole and drift-tube linac so that a good beam transmission is obtained, and this is the method adopted by AVO in their LIGHT product offering [89]; at present their RFQ is under development [90]. It is difficult to make a comparative costing but the cost of a single-room linac-based system is publicly quoted at $40M to be similar to other technology options (although it is unclear whether this may include a gantry).

The alternative to a conventional linac injection is to implement a so-called cyclinac in which a low-energy cyclotron (of somewhere between 24 and 70 MeV energy) acts as the injector [88]. The cyclotron has the advantage of being significantly simpler than the several components it replaces; the disadvantage is that its extraction frequency (of 10s of MHz) and bunch length is mismatched
to the much higher linac frequency, hence large beam losses over around 90% are to be expected with the concomitant component activation that will occur particularly in the first few linac cells. However, this may be acceptable given the rather low average currents that are required. The cyclinac approach has been developed in several studies, most notably the TULIP project in which the linac is mounted onto a gantry [91] (Figure 5). To make such a scheme size-competitive with other single-room solutions one requires a larger gradient than has yet been demonstrated, and work is underway at CERN and at the Cockcroft Institute to develop suitable structures. One approach is to utilise higher, X-band (c.11 GHz) frequency structures, and it is foreseen that gradients more than 70 MV/m may be obtained.

Despite their greater challenges, linacs (whether conventionally-injected or otherwise) may offer distinct advantages over more conventional cyclotrons and synchrotrons. Their higher natural repetition rate of at least several hundred hertz (for example, the TERA LIBO structures operate as high as 400 Hz) allows for pulse-by-pulse energy variation, which removes the need for energy selection and which allows for finer control of dose delivery. This is also potentially true when applied to carbon therapy, where again the TERA foundation have developed test structures (CABOTO) for acceleration of these ion species [92].

A possible use of cyclinacs is to provide protons both for therapy and for isotope production. Two projects have recently been put forward with this aim. In Italy the PERLA project intends to use a high-current 24 MeV cyclotron to perform proton irradiation of $^{100}$Mo targets to produce $^{99m}$Tc for use in single-photon computed emission tomography and other procedures; $^{99m}$Tc is used in over 80% of nuclear medicine worldwide and there is insecurity in its present primarily reactor-based supply, hence a number of projects are seeking to supply it [93, 94]. PERLA intends to combine its production with co-use of the same cyclotron as a cyclinac injector for proton therapy [95]. Similarly, the UK ProTec proposal seeks to use a high-current 24 MeV cyclotron to produce radioisotopes whilst also using it to provision protons for clinical research [96].
Dielectric Wall Accelerators

Dielectric wall accelerators (DWAs) are a class of particle accelerator derived from technology originally used to drive high-intensity induction accelerators for flash radiography. Development of suitable high-gradient insulators may potentially sustain accelerating electric fields in excess of 100 MV/m, wherein a dipole-pulsed Blumlein-like drive causes an accelerating pulse to co-propagate with an accelerating bunch [97, 98, 99]. This approach has been developed by CPAC, and some structures were already demonstrated in 2012 that delivered in excess of 15 MV/m; if such structures can act sequentially on a single proton bunch then it is likely that the resulting beam quality after acceleration will be comparable to that available from linacs. However, since the demonstration of those test structures little progress has been reported, and it is unclear what stage the development is at.
2.6. Laser Proton Acceleration

Laser-driven acceleration has the same motivation as DWA, that is to eventually obtain a very high accelerating gradient and sufficient overall acceleration length to obtain clinical-energy protons. In principle laser-based schemes are capable of obtaining acceleration gradients in excess of 1 GV/m, some way above the maximum gradients of around 100 MV/m which are imagined for conventional accelerating structures. There are several possible laser-based schemes and a number of general reviews are available [100, 101, 102]; the most notable current method is the use of the so-called target normal sheath acceleration (TNSA) method [103]. In TNSA protons are accelerated from the rear surface of a thin foil illuminated by a very short, intense laser pulse. The incoming laser energy is converted by several mechanisms into a beam of effectively-free electrons with large densities and moderate, relativistic energies. As those electrons exit the target rear surface they may generate a very intense electric field which draws ions from the foil. Acceleration of both protons and ions has been demonstrated in pilot experiments up to 10s of MeV, and in the single-pulse intensities are already comparable to what is needed to conduct patient irradiation. In principle expected increases over the next few years in available laser pulse intensity and repetition rate will give sufficient energy and intensity from such a source [101, 104, 23, 105].

However, there remain a number of challenges to using laser-driven proton acceleration [106]. Firstly, whilst the end-point energy of the protons is approaching a useful value the bulk of the accelerated protons have much lower energies. If used clinically these will have to be separated and dumped leading both to a low effective transmission and an accompanying large activation and production of contaminant neutrons [107]. Some discussion has been made of eventually achieving a smaller energy spread but as yet there is no demonstrated mechanism for achieving this from the laser acceleration itself. Some researchers have proposed using a linac to perform post-acceleration of the laser-generated bunches in part to reduce the final energy spread, but this seems a compromise that does not make use of the principle advantage of laser acceleration —
that is, to avoid the use of bulky conventional accelerator hardware. Another important issue with laser-accelerated protons (and carbon ions) is the likely very large pulse-to-pulse energy and beam quality variation. The raw fluence of the protons is likely to be eventually technically achievable; several Gy of dose could in principle be administered in a single pulse to the patient. However, a beam which is unstable in quality and intensity will likely require a reduction in intensity to allow suitable monitoring [41]; sufficiently high dose rates might have different RBE, although recent studies do not yet indicate a significant difference [23, 103, 109, 110, 111, 112]. As yet little serious attention has been given either to the delivery of laser-accelerated protons to patients or to how they might be scanned and monitored [109], and this is something that should be attended to if laser acceleration is to be developed as an alternative modality to conventional accelerators.

3. Recent trends and developments

The previous sections have discussed a number of technology developments, and it is worth summarising what are seen as the key recent and likely future trends. Firstly, it seems highly likely that superconducting cyclotrons will become much more widespread. A number of companies are developing high-field cyclotron solutions, and the marketplace is likely to benefit from the cost reductions possible from the accompanying reduction in physical mass of such systems. Since cyclotrons give an inherently stable, quasi-continuous output, they are favoured despite the relatively large losses which must be incurred in their energy selection systems. Reducing their cost will make it more difficult for new technology entrants such as linacs to compete, since the market appears to be presently driven by the desire to reduce the cost treatments rather than to introduce higher-capability technology [113]. The desire of smaller clinical centres to obtain proton therapy has similarly driven the market toward single-room solutions, and all the manufacturers are now offering products that offer a lower capital cost to begin proton therapy.
The desire to reduce both cost and footprint appears to be driving novel approaches in gantry development [114, 115]. Similarly to the greater adoption of superconducting cyclotrons is the burgeoning use of superconducting technology in gantries. Originally proposed as a method to reduce the size of carbon-ion gantries, a number of researchers have developed solutions based on novel magnet designs. For example, NIRS have developed a cryogen-free rotatable superconducting dipole-quadrupole magnet for their carbon-ion gantry [116, 117] (Figure 6), and similar developments are in progress at Lawrence Berkeley Laboratory aimed both at carbon ions and at protons [118]. These designs are often adaptations of the Pavlovic-style gantry optics [119] utilised for example at Paul Scherrer Institute [120] (Figure 7), HIT [68] and other locations, but another trend is the use of modular achromatic optics. One notable design here is the superconducting gantry proposed by ProNova which utilises a separated-function achromat [121]; the high field allows a large reduction in overall gantry mass (from 200 tons in a typical normal-conducting gantry to less than 25 tons in a superconducting one), whilst also giving greater patient and equipment access. In fact, the latter advantage is seen as very important. The first installation of a ProNova gantry is at the Knoxville Provision Center. Whilst needing more development, FFAG gantries may in the future offer the possibility of high-rate energy scanning, since they offer a much larger bandwidth than do traditional gantries where the magnetic fields must track changes in the energy of the transported particle beam [122].

Linac-based systems are likely to find it difficult to compete with more established accelerator technologies unless costs can be driven down, possibly through the development of high-gradient structures at higher frequencies than are commonly used today, i.e. X band. A step on the way to a complete high-gradient system could be to use a linac as a post-acceleration of a conventional cyclotron. The rationale for this is to increase the energy beyond that available from such cyclotrons to deliver suitable probe bunches for proton computed tomography and more generally as a radiographic range probe to help validate other imaging modalities during treatment [123, 124, 125, 126, 127]; but we note that at least
Figure 6: Engineering concept for the NIRS superconducting carbon therapy gantry. Ten combined-function magnets (dipole + quadrupole) provide dipole fields more than 3 T, reducing the size and mass of the gantry from the 600 tons of the normal-conducting HIT gantry to around 200 tons, similar to that of the conventional normal-conducting gantries presently used for proton therapy up to 250 MeV. (Courtesy Dr. Yoshiyuki Iwata/NIRS)

One manufacturer (ProTom) offers a synchrotron that achieves the required proton energies up to around 330 MeV [128]. The dose required for tomography is several orders of magnitude less than is needed for treatment, and hence the associated loss and activation in such a booster linac will be much less than if such structures were used for treatment. Both Paul Scherrer Institute and The Christie Hospital have proposals to construct such a booster linac [52, 96], which will be needed in addition to the radiographic instruments currently under development at various centres if proton tomography is to be carried out on the bodies of adult patients.

The alternative to using a booster linac for tomography is to use a single accelerator both for treatment and for imaging. Here the most likely solution

20
Figure 7: Installation of the Paul Scherrer Institute Gantry 2, which utilises scanning magnets upstream of the final dipole in a Pavlovic-style optics also used at HIT but no commonly in other centres. Parallel beam scanning is enabled at the expense of a larger, more complex final bending magnet.

is the proton FFAG, and solutions have been proposed both at Fermilab and in the UK [129, 82]. In both cases the trade-off is to have an accelerator capable of higher-energy output but which is necessarily larger and more complex. Given the present market it is likely that the first example of such a machine would have to be constructed at an accelerator research centre.

Whilst market offerings for charged particle therapy are dominated by protons, carbon-ion systems may also be obtained from several companies. The first such centres were obtained mainly as collaborations with accelerator laboratories such as the HIT synchrotron designed in collaboration with GSI in Germany, and the European PIMMS study which resulted in the two treatment centres at CNAO and MedAustron. More recently several companies have offered complete carbon-ion systems; these include the Siemens centres at Marburg and Shanghai, and several recent projects in Japan and other eastern countries involving Hitachi, Mitsubishi, Sumitomo and Toshiba [130]. In addition there are proposals for carbon-ion therapy facilities across Europe and the USA. The great advances in charged particle therapy technology in recent years — most significantly the widespread adoption of spot scanning pioneered at Paul Scherrer Institute [131, 132] — mean that proton therapy is now perceived as an
established technology; attention is now increasingly turning to improving the use of other ions. Whilst other more exotic ions and species such as pions \[133\] and antiprotons \[134, 135, 136\] were examined in the past, carbon remains the favoured ion species in use today and thus developments focus there. Most heavy-ion proposals involve the provision of protons and carbon ions from the same accelerator facility, as exemplified by the PIMMS and Siemens designs. The UK-led PAMELA study examined a combination FFAG facility that also delivered both protons and carbon ions, but at a faster bunch rate (1 kHz) than might be achieved from synchrotrons; future designs such as this will need to reduce cost mostly likely from size reduction of the accelerator, for example by adopting higher fields needing superconductivity.

Following on from some early work in Japan, a likely future direction for carbon-ion therapy is the use of either carbon-11 or carbon-10 ions both for treatment and for simultaneous range verification; range verification is seen as an important next step in the improvement generally of charged particle therapy \[114, 137, 138\]. Carbon-11 is already quite widely used directly for positron-emission tomography, despite its short half life; treatment can be performed with carbon-11 ions, which when stopped in the patient will give rise to a natural range signature when they beta decay. Use of carbon-11 ions requires their manufacture in a suitable manner, and this has been examined using several methods which may be broadly grouped either into batch production (the carbon-11 is manufactured and then physically transported to be used in an ion source), or online production where one accelerator is used to manufacture carbon-11 prior to its direct injection to the therapy machine. Synchrotrons require relatively little adaptation to change from being used for carbon-12 to carbon-11, and perhaps carbon-10 may also be used.

Unstable carbon ions may be used they will still require the large infrastructure inherent for such large-rigidity species. Another avenue is to examine the use of helium ions as a ‘half-way house’ between protons and carbon ions. Whilst it is difficult to construct a compact carbon-ion facility it may be possible to construct one for helium treatment; helium may have some advantages in
terms of lateral scatter [139]. However, very little has yet been done to examine the role of helium ions in modern charged particle therapy [140, 141], and a study of such a machine is probably warranted.

4. Summary

The discussion above has attempted to describe the most notable recent developments and remaining challenges in accelerator technology applied to charged particle therapy. Whilst a number of technologies have been developed, the cyclotron remains the workhorse of proton therapy today but faces some challenges if competing technologies such as laser acceleration, FFAGs or high-gradient linacs can be made capable and cheap enough. Superconducting technology developments are the key trend here. In carbon treatment the synchrotron is likely to remain the primary accelerator used, and the most likely avenues for significant development will be the adoption of superconducting magnets in the gantries and accelerator sources, and the eventual use of carbon-11 and carbon-10 ions.

5. Acknowledgements

The authors would like to thank their colleagues at Manchester, Christie, UCLH and Paul Scherrer Institute for useful discussions. Also thanks in particular to Jose Alonso and Jay Flanz for their earlier reviews of this topic. This work was supported in part by grants from the Science and Technology Facilities Council (Cockcroft Institute Core Grant No. ST/G008248/1), and from the European Union (FP7 network EUCARD2).

References


[34] O. Jaekel, A. Smith, C. Orton, The more important heavy charged particle radiotherapy of the future is more likely to be with heavy ions rather than protons, Medical Physics 40 (2013) 090601.


R. Wideröe, Über ein neues Prinzip zur Herstellung hoher Spannungen, Archiv für Elektrotechnik 21 (4) (1928) 387.


A. Garonna, et al., Cyclinac medical accelerators using pulsed C\textsuperscript{6+}/H\textsubscript{2}\textsuperscript{+} ion sources, Journal of Instrumentation 5 (2010) C09004.


